

State space modeling approach to decompose daily sales of a restaurant into time-dependent multi-factors

Yamaguchi, Rui
Faculty of Mathematics, Kyushu University

Tsuchiya, Eiko
School of Science, Tokyo Institute of Technology

Higuchi, Tomoyuki
Institute of Statistical Mathematics

<http://hdl.handle.net/2324/3346>

出版情報 : MHF Preprint Series. MHF2004-2, 2004-03-03. 九州大学大学院数理学研究院
バージョン :
権利関係 :



MHF Preprint Series

Kyushu University
21st Century COE Program
Development of Dynamic Mathematics with
High Functionality

State space modeling approach
to decompose daily sales of a
restaurant into time-dependent
multi-factors

R. Yamaguchi, E. Tsuchiya
T. Higuchi

MHF 2004-2

(Received March 3, 2004)

Faculty of Mathematics
Kyushu University
Fukuoka, JAPAN

State space modeling approach to decompose daily sales of a restaurant into time-dependent multi-factors

Rui Yamaguchi ^{a,*} Eiko Tsuchiya ^b Tomoyuki Higuchi ^c

^a*Faculty of Mathematics, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka, Japan*

^b*School of Science, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo, Japan*

^c*Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo, Japan*

Abstract

Daily sales of catering establishments are affected by several factors: The day of the week, national holidays, weather, events taking place near the site, etc. Therefore, it is beneficial to construct a model decomposing a time series of the sales into such factors. Thus, the model will allow us to predict future sales with good accuracy in order to plan various management strategies such as purchasing, staff assignments, new store openings, etc. In this study, we propose a prediction method for the daily sales of a catering establishment based on a state space model framework in which an optimal model is selected by an information criterion; we then apply it to 2-year daily sales data of a restaurant that is adjacent to a large-scale convention center and office buildings.

Key words: state space model, marketing, knowledge discovery

1 Introduction

Since the collapse of Japan's so-called "bubble economy" in the 1990s, the economy has been engulfed in waves of depression. Thus, all industries are

* R. Yamaguchi is a corresponding person.

Email addresses: ruiy@math.kyushu-u.ac.jp (Rui Yamaguchi), eicot0808@hotmail.com (Eiko Tsuchiya), higuchi@ism.ac.jp (Tomoyuki Higuchi).

demanding efficiency in their business and continuing to seek more efficient and effective marketing strategies. Due to the maturation of markets and increase in the variety of customers' preferences, they have shifted emphasis from conventional mass-marketing aimed at the general public to micro-marketing aimed at more specific small segments, moreover, to one-to-one marketing targeting individuals (Montgomery, 1997; Peppers and Rogers, 1997; Lilian and Rangasamy, 1997, 2002).

For catering industries, due to maturation of the markets and the excessive number of stores, even restaurant chains and fast-food establishments, etc., have begun activities to pull themselves out of a uniform administration and differentiate themselves from others. To respond to the diversification of consumer preferences and changes of constituencies, business managers have adopted strategies differing from the conventional path, such as offering a limited menu in the area, and using service uniforms and interiors to create a unique and individual restaurant atmosphere (Japan Food Service Association, 2003). In other words, they manage such restaurants by considering the situation inherent in each of them.

It is important for a catering establishment to predict both a long-term variation of sales (trend) and daily sales. Since the stock in hand is food, which spoils rapidly, daily sales in particular require more precise prediction than those for other industries. Daily sales of catering establishments are affected by several factors: The day of the week, national holidays, weather, events taking place near the site, etc. Therefore, it is beneficial to construct a model decomposing a time series of the sales into such factors. Kondo and Kitagawa (2000); Kitagawa et al. (2003) analyzed daily scanner sales data in a store using models that decompose the time series into trend, day-of-the-week effect, and explanatory variable effect due to price promotion. Such a model will allow us to predict future sales with good accuracy in order to plan various management strategies such as purchasing, staff assignments, new store openings, etc. In this study, we propose a prediction method for daily sales of a catering establishment based on a state space model framework in which an optimal model is selected by an information criterion; we then applied it to 2-year daily sales data of a restaurant that is adjacent to a large-scale convention center and office buildings.

The organization of the paper is as follows. In Section 2, we propose the model. In Section 3, we explain the estimation of a state vector and optimal parameters. In Section 4, we show the results obtained by applying the model to the data and discuss them. In Section 5, we evaluate the predictive ability of the model and discuss it. Section 6 contains our conclusion.

2 Model

In this study, we used three types of time series data for the sales of the restaurant at time n , that is, $y_{1,n}$, $y_{2,n}$, and $y_{3,n}$, which represent daily sales for lunches, for drinking parties, and for the total, respectively. The time n is an index indicating the number of days starting from January 1, 2000. We also used time series data: $X_{1,n}$, $X_{2,n}$, and d_n , where $X_{1,n}$ is an anticipated mean attendance figure per day for events taking place at the large-scale convention center, $X_{2,n}$ and d_n are weather and day of the week (see Table1). The source of the time series data of the sales $y_{i,n}$, ($i = 1, 2, 3$) and weather $X_{2,n}$ is daily sales reports written by the store manager of the restaurant. The anticipated mean attendance figure $X_{1,n}$ is based on the anticipated cumulative attendance figure during a session of an event described in a monthly event schedule published by a proprietary company of the convention center, which will be detailed in Section 2.3.

To explain the three types of sales data $y_{i,n}$, ($i = 1, 2, 3$), we constructed a model for each data set to decompose them into the above mentioned factors. In the following, the analyzed sales data are represented by y_n . The decomposing model we constructed is as follows:

$$y_n = t_n + W_n + R_n + E_n + r_n + \varepsilon_n, \quad (1)$$

where t_n , W_n , R_n , E_n , r_n and ε_n are the trend component, the weekly effect, the rain effect, the event effect, the auto-regression (AR) model component, and the residual component, respectively. In this model, we assume that t_n follows a second order trend model given by $t_n = 2t_{n-1} - t_{n-2} + v_{t,n}$, $v_{t,n} \sim N(0, \tau_t^2)$ (Kitagawa and Gersch, 1996). We also assume that r_n follows a second order stationary AR model given by $r_n = \sum_{j=1}^2 a_j r_{n-j} + v_{r,n}$, $v_{r,n} \sim N(0, \tau_r^2)$. There is no specific reason to set it as the second order and it is possible to change the order as the need arises. In the rest of this section, W_n , R_n , and E_n are explained in detail.

2.1 Model: Weekly Effect

We assume that the weekly effect W_n can be expressed in the following form:

$$W_n = w_n + h_{1n}\beta_1(w_{Sun,n} - w_n) + h_{2n}\{\beta_2(w_{Fri,n} - w_n) + \beta_3(w_{Sat,n} - w_n)\}. \quad (2)$$

The first term w_n is the day-of-the-week effect, which is expressed by a seasonal

adjustment model with a period of seven:

$$\sum_{j=0}^6 w_{n-j} = v_{w,n}, \quad v_{w,n} \sim N(0, \tau_w^2) \quad (3)$$

(Kitagawa and Gersch, 1984, 1996; Harvey, 1989). This term accounts for a basic pattern in a week.

The second term represents an effect when a current day is a national holiday in a one-week period from Monday to Friday, which we call the “holiday effect,” where $h_{1n} \in \{0, 1\}$ is an indicator function of the term (see Table 2). We note that in Japan, there are fourteen national holidays in a year. Thus, there is approximately one national holiday per month. $\beta_1(w_{Sun,n} - w_n)$ in the term is the model of our expectation, that is, sales on a national holiday in a one-week period from Monday to Friday probably resemble those of a Sunday. This is represented by the product of the difference between a day-of-the-week effect of the most recent Sunday of a current day ($w_{Sun,n}$) and that of the current day (w_n), and a coefficient of the similarity ($0 \leq \beta_1 \leq 1$). In this model, if $\beta_1 = 0$, the weekly effect of a day with a holiday effect becomes $W_n = w_n + 0$, that is, represented only by a day-of-the-week effect of the current day. Whereas if $\beta_1 = 1$, a weekly effect of a day with a holiday effect becomes $W_n = w_n + (w_{Sun,n} - w_n) = w_{Sun,n}$, that is, represented only by a day-of-the-week effect of the most recent Sunday of the current day.

The third term represents an effect of the preceding day of a national holiday, which we call the “holiday-eve effect,” where $h_{2n} \in \{0, 1\}$ is an indicator function (see Table 2). $\{\beta_2(w_{Fri,n} - w_n) + \beta_3(w_{Sat,n} - w_n)\}$ is also a model of our expectation for sales on the day preceding a national holiday, that is, the sales are similar to those of Friday or Saturday. This is also represented by the sum of two products where $w_{Fri,n}$ and $w_{Sat,n}$ are day-of-the-week effects of the most recent Friday and Saturday of a current day, and β_2 and β_3 are coefficients of their similarities ($0 \leq \beta_1, \beta_2 \leq 1$) under a binding condition: $\beta_2 + \beta_3 \leq 1$.

2.2 Model: Rain Effect

The rain effect R_n represents the effect of weather on sales. We assume that R_n is given by,

$$R_n = \gamma_R f_R(X_{2,n}), \quad (4)$$

where γ_R is a time-invariant coefficient, which is in units of money and $f_R(X_{2,n})$ is a function of a categorized weather $X_{2,n} \in \{1, 2, 3, 4, 5\}$ (fine, cloudy, rain, heavy rain, and snow), as seen in Figure 1. The function $f_R(X_{2,n})$ in the rain effect $R_n = \gamma_R f_R(X_{2,n})$ takes values within $0 \leq f_R(X_{2,n}) \leq 1$ (see Figure 1).

Thus the meaning of the coefficient γ_R included in the state vector changes with its sign. If it takes a positive value, the restaurant attracts customers in bad weather. On the other hand, if it takes a negative value, the restaurant loses customers in bad weather. There is no specific reason for the form of this function, it was constructed based on various experiences.

2.3 Model: Event Effect

The event effect E_n is a model to estimate effects from attendance figures for events taking place at the large-scale convention center. We assume the model as follows:

$$E_n = \alpha_{B,n} f_B(X_{1,n}) + I(f_B(X_{1,n}))(\alpha_w f_w(d_n) + h_{3n} \alpha_h). \quad (5)$$

The first term is a basic term when there are events, which is represented by the product of a non-linear function $f_B(X_{1,n})$ and a time-variant coefficient $\alpha_{B,n} = \alpha_{B,n-1} + v_B$, $v_B \sim N(0, \tau_B^2)$. $f_B(X_{1,n})$ is modeled to show the effect that emerges when $X_{1,n}$ is over a threshold value X_{Eve} , and has a plateau after the number of visitors reaches the capacity of the restaurant; this is given as,

$$f_B(X_{1,n}) = \begin{cases} \frac{g(X_{1,n})}{\max\{g(X_{1,n})\}} & X_{1,n} \geq X_{Eve} \\ 0 & X_{1,n} < X_{Eve} \end{cases} \quad (6)$$

where

$$g(X_{1,n}) = 1 - \exp\{-a(X_{1,n} - X_{Eve})\} \quad (7)$$

(see Figure 2).

We note that the anticipated mean attendance figure $X_{1,n}$ does not represent differences in attendance figures due to variations of the day of the week and national holidays. $X_{1,n}$ is simply obtained as an average of an anticipated cumulative attendance figure during a session of an event because the published monthly event schedule only has anticipated cumulative attendance figures. The differences are represented in the second term.

$I(\cdot)$ in the second term is a step function whereby the second term takes a value only when there is an event effect. There are two terms in a parenthesis term after $I(f_B(X_{1,n}))$. The first term in the parentheses, which we call “the first correction term,” is a term that represents variations of attendance figures for each day of the week, where $f_w(d_n)$ is a function of the day of the week $d_n \in \{1, 2, \dots, 7\}$ (*Sun, Mon, \dots, Sat*), and α_w is a time-invariant coefficient. The second term in the parentheses, which we call “the second correction term,” is for an augmentation of sales due to an event on holidays (i.e., Saturday,

Sunday, and a national holiday), where α_h is a time-invariant coefficient in units of money, and $h_{3n} \in \{0, 1\}$ is an indicator function (see Table 2).

Here we note that the correction for Saturday and Sunday is made redundantly by the first and the second correction term. This is because we have made the model step by step, which will be explained in more detail in Section 4.4. First, we implemented the augmentation term for holidays α_h (the second correction term), because it is natural to expect that attendance for events during holidays will increase compared to that during weekdays. However, it is difficult to estimate the above mentioned $f_w(d_n)$ simultaneously. Thus, we first estimated prediction errors using a model without the first correction term, then obtained averages for them for each day of the week. We considered these average values as a mean variation pattern of sales due to variations in attendance figures for each day of the week; then we made the function $f_w(d_n)$ from them. The first correction term was made as a product of $f_w(d_n)$ and a constant coefficient α_w . As a result, the correction for Saturday and Sunday was made by both the first and second correction terms in this model.

2.4 Model: State Space Representation

The model described in the previous sections can be represented by a state space model (Harrison and Stevens, 1976; Anderson and Moore, 1979; Kitagawa and Gersch, 1984, 1996):

$$x_n = F_n x_{n-1} + G_n v_n \quad \text{[System Model]} \quad (8)$$

$$y_n = H_n x_n + \varepsilon_n \quad \text{[Observation Model]} \quad (9)$$

by setting the state space vector as

$$x_n = \left[x_{t,n}^t \mid x_{W,n}^t \mid x_{R,n} \mid x_{E,n}^t \mid x_{r,n}^t \right]^t, \quad (10)$$

where $x_{t,n}$, $x_{W,n}$, $x_{R,n}$, $x_{E,n}$, and $x_{r,n}$ are partial vectors relating to the trend, the weekly effect, the rain effect, the event effect, and the AR-model component respectively. Each of these partial vectors is given as follows:

$$x_{t,n} = [t_n, t_{n-1}]^t \quad (11)$$

$$x_{W,n} = [w_n, w_{n-1}, w_{n-2}, w_{n-3}, w_{n-4}, w_{n-5}]^t \quad (12)$$

$$x_{R,n} = [\gamma_R] \quad (13)$$

$$x_{E,n} = [\alpha_{B,n}, \alpha_w, \alpha_h]^t \quad (14)$$

$$x_{r,n} = [r_n, r_{n-1}]^t. \quad (15)$$

We note that $\gamma_R, \alpha_w, \alpha_h$ in the state vector become time-invariant as a result of smoothing, since they do not have system noises.

The state transition matrix F_n is a 14×14 matrix given by

$$F_n = F = \begin{bmatrix} F_t & & & & \\ & F_W & & & \\ & & F_R & & \\ & & & F_E & \\ & & & & F_r, \end{bmatrix} \quad (16)$$

where F_t , F_W , F_R , F_E , and F_r are partial matrices relating to the trend, the weekly effect, the rain effect, the event effect, and the AR-model component, respectively. Each of the partial matrices is given by,

$$F_t = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \quad (17)$$

$$F_W = \begin{bmatrix} -1 & -1 & \dots & -1 \\ 1 & & & \\ & \ddots & & \\ & & 1 & 0 \end{bmatrix} \quad (18)$$

$$F_R = \begin{bmatrix} 1 \end{bmatrix} \quad (19)$$

$$F_E = \begin{bmatrix} 1 \\ & 1 \\ & & 1 \end{bmatrix} \quad (20)$$

$$F_r = \begin{bmatrix} a_1 & a_2 \\ 1 & 0 \end{bmatrix}. \quad (21)$$

The observation matrix H_n is a 14-dimensional row vector represented by

$$H_n = [H_{t,n} \mid H_{W,n} \mid H_{R,n} \mid H_{E,n} \mid H_{r,n}], \quad (22)$$

where $H_{t,n}$, $H_{W,n}$, $H_{R,n}$, $H_{E,n}$, $H_{r,n}$ are partial row vectors relating to the trend, the weekly effect, the rain effect, the event effect, and the AR-model component, respectively.

The partial observation matrices for the trend $H_{t,n}$, the rain effect $H_{R,n}$, and

the AR-model component $H_{r,n}$ are represented as follows:

$$H_{t,n} = H_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad (23)$$

$$H_{R,n} = \begin{bmatrix} f_R(X_{2,n}) \end{bmatrix} \quad (24)$$

$$H_{r,n} = H_r = \begin{bmatrix} 1 & 0 \end{bmatrix} . \quad (25)$$

The partial observation matrix for the event effect is given by,

$$H_{E,n} = \begin{bmatrix} \zeta, & I(\zeta)f_w(d_n), & I(\zeta)h_{3n} \end{bmatrix} , \quad (26)$$

where $\zeta = f_B(X_{1,n})$.

The partial observation matrix for the weekly effect changes its form according to the following three cases switched by indicator functions $h_{1,n}, h_{2,n} \in \{0, 1\}$ for a current day (see Table 2).

- (1) A day, which is neither a national holiday nor a holiday eve: $(h_{1,n}, h_{2,n}) = (0, 0)$

In this case, the weekly effect becomes $W_n = w_n$ by Eq. (2). Then the partial state space vector for the weekly effect $x_{W,n}$ (Eq. (12)) and Eqs. (1) and (9) yield

$$W_n = H_{W,n}x_{W,n} . \quad (27)$$

Thus, $H_{W,n}$ is represented by,

$$H_{W,n} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} . \quad (28)$$

- (2) A national holiday: $(h_{1,n}, h_{2,n}) = (1, 0)$

In this case, $W_n = (1 - \beta_1)w_n + \beta_1w_{Sun,n}$ by Eq. (2). Here the day-of-the-week effect of the most recent Sunday of the current day ($w_{Sun,n}$) is included in $x_{W,n}$. The position in $x_{W,n}$ is variable according to the day of the week of the current day (d_n), for example, $w_{Sun,n} = w_{n-1}$ when the current day is Monday ($d_n = 2$). Thus, according to Eq. (27), $H_{W,n}$ is given by,

$$H_{W,n} = \begin{cases} [\kappa \ \beta_1 \ 0 \ 0 \ 0 \ 0], & \text{Mon } (d_n = 2) \\ [\kappa \ 0 \ \beta_1 \ 0 \ 0 \ 0], & \text{Tue } (d_n = 3) \\ [\kappa \ 0 \ 0 \ \beta_1 \ 0 \ 0], & \text{Wed } (d_n = 4) \\ [\kappa \ 0 \ 0 \ 0 \ \beta_1 \ 0], & \text{Thu } (d_n = 5) \\ [\kappa \ 0 \ 0 \ 0 \ 0 \ \beta_1], & \text{Fri } (d_n = 6) \end{cases} , \quad (29)$$

where $\kappa = 1 - \beta_1$.

(3) A holiday eve: $(h_{1n}, h_{2n}) = (0, 1)$

In this case, $W_n = (1 - \beta_2 - \beta_3)w_n + \beta_2 w_{Fri,n} + \beta_3 w_{Sat,n}$ by Eq. (2). As is similar in the case of the holiday, the day-of-the-week effects of the most recent Friday and Saturday ($w_{Fri,n}$ and $w_{Sat,n}$) included in $X_{W,n}$ also change positions according to the day of the week of the current day, for example, $w_{Fri,n} = w_{n-3}$ and $w_{Sat,n} = w_{n-2}$ when the current day is Monday ($d_n = 2$). However, when the current day is Thursday ($d_n = 5$), the day-of-the-week effect of the most recent Friday is given by $w_{Fri,n} = w_{n-6}$, which is not explicitly included in $x_{W,n}$. However, in this case, by using components of $x_{W,n}$ following the seasonal adjustment model of Eq. (3), it can be represented as $w_{Fri,n} = -\sum_{j=0}^5 w_{n-j}$. Thus, according to Eq. (27), $H_{W,n}$ is given by,

$$H_{W,n} = \begin{cases} [\xi \ 0 \ \beta_3 \ \beta_2 \ 0 \ 0], & \text{Mon } (d_n = 2) \\ [\xi \ 0 \ 0 \ \beta_3 \ \beta_2 \ 0], & \text{Tue } (d_n = 3) \\ [\xi \ 0 \ 0 \ 0 \ \beta_3 \ \beta_2], & \text{Wed } (d_n = 4) \\ [\chi \ \psi \ \psi \ \psi \ \psi \ \omega], & \text{Thu } (d_n = 5) \end{cases}, \quad (30)$$

where $\xi = 1 - \beta_2 - \beta_3$, $\chi = 1 - 2\beta_2 - \beta_3$, $\psi = -\beta_2$, and $\omega = \beta_3 - \beta_2$.

3 State Vector and Parameter Estimation

Since the model described in the previous section can be expressed in a state space model (Eqs. (8) and (9)), we can estimate the state and decompose the time series efficiently by employing the Kalman Filter and the fixed interval smoother algorithms (Kalman, 1960; Kitagawa and Gersch, 1984, 1996; Harvey, 1989; West and Harrison, 1997). We denote a conditional mean and a conditional variance-covariance matrix of a state vector x_n with a given time series $Y_j = \{y_1, \dots, y_j\}$ respectively as follows:

$$x_{n|j} \equiv E(x_n | Y_j) \quad (31)$$

$$V_{n|j} \equiv E[(x_n - x_{n|j})(x_n - x_{n|j})^t]. \quad (32)$$

For a state estimation, the state at a time n is classified into the following three types corresponding to contexts of the time n and a time j of a time series Y_j used for the estimation, that is, “prediction” when $j < n$, “filtering” when $j = n$, and “smoothing” when $j > n$.

The goodness of a statistical model can be evaluated by the predictive ability. The likelihood of a model is a value approximately evaluating the predictive ability of the model. Thus, by regulating values of parameters included in a model to maximize the likelihood, we can approximately maximize the predictive ability of the same model. In general, the likelihood of a time series

model is given by,

$$L(\theta) = p(y_1, \dots, y_N | \theta) = \prod_{n=1}^N p(y_n | Y_{n-1}, \theta), \quad (33)$$

where each component $p(y_n | Y_{n-1}, \theta)$ can be obtained as a byproduct of the Kalman filter (Kalman, 1960). If there are several candidate models, the goodness of the fit of the models can be evaluated by the AIC criterion defined by

$$\text{AIC} = -2 \log L(\hat{\theta}_{ML}) + 2(\text{number of parameters}) \quad (34)$$

(Akaike, 1974; Sakamoto et al., 1986), where $\hat{\theta}_{ML}$ is the maximum likelihood estimate on θ .

The model used in this study includes an unknown parameter vector: $\theta = [\sigma^2, \tau_t^2, \tau_w^2, \tau_B^2, \tau_r^2, \beta_1, \beta_2, \beta_3, a_1, a_2]^t$. In the following section, we will demonstrate the result of a decomposition of the sales data for lunch, which is obtained by applying the above-mentioned model. For the actual optimization procedure, we used a parameter vector θ^\dagger made from a partial vector of θ in which σ^2 has been removed and the other variance parameters (τ_t^2 , τ_w^2 , τ_B^2 , and τ_r^2) have been divided by σ^2 , that is, $\theta^\dagger = [\tau_t^2/\sigma^2, \tau_w^2/\sigma^2, \tau_B^2/\sigma^2, \tau_r^2/\sigma^2, \beta_1, \beta_2, \beta_3, a_1, a_2]^t$. We fixed the ranges for each of the variance parameters to be searched by the grid search procedure as $-5 \leq \log_{10}(\tau_t^2/\sigma^2) \leq -4$, $-5 \leq \log_{10}(\tau_w^2/\sigma^2) \leq -4$, $-4 \leq \log_{10}(\tau_B^2/\sigma^2) \leq 0$, and $-4 \leq \log_{10}(\tau_r^2/\sigma^2) \leq 0$, where the width between the grids was set as 0.5. The width between the grids for β_1 , β_2 , and β_3 was set as 0.2. The AR coefficients a_1 and a_2 were set to satisfy a stationary condition, for which the range of partial-autocorrelation coefficients (PARCOR) was limited between -1 and 1. For the actual procedure, we set the absolute values of the PARCOR to less than 0.95 to avoid a compensation with the trend component. Notably, the estimated state variables shown in the following section were made by smoothing ($x_{n|N}$ and $V_{n|N}$ where $N = 731$) unless referred to otherwise.

4 Results of the Analysis

4.1 Result: Trend

Figure 4 shows the original time series of the sales of lunches (y_n : Thin line) and the estimated trend component (t_n : Thick line). In Figure 4, we can see that the trend represents the long-term variation in the original data. We can also see that even trends for missing parts in the original data have been estimated by the Kalman filter and the fixed interval smoother. In Figure 5, year-round variations of the trend for each year have been superposed. There

are similar seasonal variations for both of them. These variations occurring around the same time in both years show that the trend of the year 2001 (solid line) is placed over that of the year 2000 (dashed line) except for a short interval at the end of both years. Thus, restaurant was evidently getting a boost in sales during the period in general.

4.2 Result: Weekly Component

The weekly component W_n (see Eq. (2)) is shown as follows. Figure 6 is a plot made by superposing the day-of-the-week effect w_n in a week $d_n \in \{1 \dots 7\}$ for the whole period. We can recognize a pattern in the plot, that is, a cutdown in sales during the weekend and a pickup in sales during weekdays.

By considering the fact that the restaurant is adjacent to office buildings, we can surmise from this result that the basic pattern of the lunch sales in a week corresponds to the movements of customers from the office buildings; they work on weekdays and rest on weekends. Figure 7 shows the holiday effect $h_{1n}\beta_1(w_{Sun,n} - w_n)$, (solid line) and the holiday-eve effect $h_{2n}\{\beta_2(w_{Fri,n} - w_n) + \beta_3(w_{Sat,n} - w_n)\}$, (dashed line). For this data, the optimal values of β_1 , β_2 , and β_3 are 1.0, 0, and 0.4, respectively. This means a weekly effect with a holiday effect can be expressed by $W_n^{Holiday} = w_{Sun,n}$ due to Eq. (2), that is, it can be explained by the-day-of-the-week effect of the most recent Sunday. In the same way, a weekly effect with a holiday-eve effect can be represented by $W_n^{Holiday-Eve} = 0.6w_n + 0.4w_{Sat,n}$ due to Eq. (2), which means the weekly effect of the holiday-eve is in part similar to that of the most recent Saturday, but dissimilar to that of the most recent Friday. As shown in Figure 6, $w_{Sun,n}$ and $w_{Sat,n}$ take negative values. Thus, the weekly effect of days with the holiday effect or the holiday-eve effect takes a smaller value than the weekly effect of the other day.

4.3 Result: Rain Effect

As a result of the smoothing, the constant coefficient takes a positive value: $\gamma_R = 11.3013$. Thus, as shown in Figure 8, the rain effect R_n takes positive values, which means the restaurant attracts customers in bad weather as mentioned previously in Section 2.2. Considering that the restaurant is connected to the office buildings and the convention center by arcades, this result may reflect the mentality of customers who tend to have lunch at a place nearby when the weather is bad.

4.4 Result: Event Effect

Figure 10 shows (a) the first term $\alpha_{B,n}f_B(X_{1,n})$ and (b) its coefficient $\alpha_{B,n}$ in the event effect $E_n = \alpha_{B,n}f_B(X_{1,n}) + I(f_B(X_{1,n}))(\alpha_w f_w(d_n) + h_{3n}\alpha_h)$. As mentioned previously in Section 2, the first term represents the basic part of the event effect. The value is larger than that of the weekly effect and the rain effect, and, thus explains the larger amount of sales. From these results we can confirm the bilateral character of the restaurant, that is, one aspect is the underlying character of a restaurant in a business park and the other aspect is of a restaurant adjacent to a convention center largely affected by an external variable, namely, turnouts for events.

The term $\alpha_w f_w(d_n)$ in the second term of the event effect is the correction term as a function of the day of the week ($d_n \in \{1, \dots, 7\}$) as mentioned in Section 2.3. The function $f_w(d_n)$ was yielded by the following procedure. First, we constructed a model without the correction term from Eq. (1) as shown below:

$$y_n = t_n + W_n + R_n + E'_n + r_n + e'_n, \quad (35)$$

where $E'_n = \alpha_{B,n}f_B(X_{1,n}) + I(f_B(X_{1,n}))h_{3n}\alpha_h$. Then, we calculated the time series of prediction errors $e'_{n|n-1}$. Finally, the correction function was obtained from biases in distributions of $e'_{n|n-1}$ for each day of the week. To be more precise, each of the aforementioned biases is a value made by taking the average of the time series extracted from the prediction errors $e'_{n|n-1}$ for each day of the week under the condition that the event effect existed ($I(f_B(X_{1,n})) = 1$). The correction function $f_w(d_n)$ takes a value from the biases for the corresponding day of the week d_n (see Figure 9). After obtaining $f_w(d_n)$, we constructed the model for Eq. (1) and then estimated again the whole state vector of Eq. (10). The second term α_h represents the augmentation of the sales on holidays (Saturday, Sunday, and national holidays) with events. As a result of the smoothing, it has become a positive constant ($\alpha_h = 11.7293$). It is an acceptable result from the perspective of common-sense whereby the sales on the holidays with events increase in comparison with those on weekdays with events.

4.5 Result: AR Component

Figure 11 shows r_n the stationary second-order AR model component. This term is expected to represent short-term variations. Since it shows large variations accompanied by large values of the event effect, it may also compensate for part of the sales which cannot be expressed by the event effect.

5 Discussion

The predictive ability of the model can be evaluated by a distribution of the prediction-error component $\varepsilon_{n|n-1} = y_n - H_n x_{n|n-1}$. Figure 12 (a) and (b) show the frequency of the prediction error for days without the event effect ($I(f_B(X_{1,n})) = 0$) and that for days with the event effect ($I(f_B(X_{1,n})) = 1$), respectively. For days without event effects, the prediction error is restricted to 25.699 [1000 Yen] for the standard deviation. Whereas, for days with event effects, the prediction error becomes much larger, that is, 62.673 [1000 Yen] for the standard deviation.

From this result, we can consider improving the model for the event effect in order to raise the predictive ability of the model. By referring to Eq. (5), the basic term of the event effect is characterized by an anticipated mean attendance figure for events ($X_{1,n}$). In that model, types of attendance and difference in their behavior for each type of event were not dealt with. However, it should be noted that the attendance behavior actually diverges because of the type of event and the relation between the position of the restaurant and that of the venue in the convention center. As examples of events on days with large prediction errors, we can take an event targeted at children and a certain type of product fair. Since the restaurant is targeted at adults, it is difficult for parents and children to enter. Thus, in the former case, the sales growth is much smaller than that we would expect from the anticipated attendance figure. For the latter case, such as a product fair dealing with food, there are usually sampling corners and booths selling lunch boxes. Therefore, the visitors did not go to the restaurant, which is closer to the office buildings than the convention center, accounting for the smaller amount of sales than expected. Considering the above discussion, we can develop a strategy to improve the model such as the use of different correction functions for each type of event. For an extreme case, it might be better to use sales for each event in the previous year, the so-called “case-based” method. However, in that case, the model is completely specialized only for the restaurant concerned. Therefore, the model cannot be generally applied to other restaurants. Such a generality is important formulating a plan to develop a new restaurant. Thus, we need to construct a model that can express effects from different types of events while maintaining generality. As a concrete strategy, we consider the classifications of events by contents, times, and venues, and the switching of models to be applied according to the class of events. This is an issue for future study.

The aforementioned improvement of the model was clarified only after detailed investigation of the error term, that is, effects which cannot be explained by the model. In other words, we were able to infer the movement of visitors for events by discussing the error terms. This extracted information and knowledge can be used to improve the model through the evaluation of an information cri-

terion. The newly improved model can also be used to extract further new information and knowledge and then help to construct a new model again. By continuing the above procedure, we can expect an everlasting improvement of the model and extraction of new knowledge (Higuchi and Kitagawa, 2000).

6 Conclusion

In this study, we proposed a model to decompose daily sales of a restaurant into effects from the day of the week, weather, events near the site, etc., which is based on a state space model framework. We applied the model to the real data of 2-year daily sales of a restaurant, which is adjacent to a large-scale convention center and office buildings. As a result, we obtained the trend, the weekly effect, the rain effect, the event effect, the AR-model component, and the residual. Detailed investigations of each component allowed us to infer underlying rational reasons for each of them. Even effects estimated only by so-called “hunches” at the work front (e.g. holiday effect) can be included in the model in a positive and flexible manner, and possibly inferred in a quantitative and relative manner within a framework of a state space representation and evaluation by an information criterion. We discussed the possibility of an everlasting improvement of the model and knowledge discoveries due to an analysis of residuals within the framework.

As an implementation of the presented model for a single individual restaurant, we can suggest an add-in software for the spreadsheet software. By inputting daily sales, weather forecast, etc. into the software, tomorrow’s sales would be predicted dynamically. The result can be applied to devising a strategy for laying in stock. Knowledge obtained from the estimated terms (e.g. pattern of weekly effect) is inherent information based on real data. Thus, it is useful to realize marketing, which has clear-cut target segments, that is, micro-marketing rather than mass-marketing, by using it for advertisement planning, etc. Meanwhile, for franchise chains, we suggest a Web-based software as an implementation. By employing such software, a manager in a franchiser can easily evaluate the business performances of a franchisee by comparing trends of sales for each franchisee. It may also be possible to simulate sales of a new restaurant with a different locational conditions to those of an existing restaurant. We hope to start work on plans to realize the above implementations soon.

Acknowledgment

This work is partially supported by Kyushu University 21st Century COE Program, Development of Dynamic Mathematics with High Functionality, of the Ministry of Education, Culture, Sports, Science and Technology of Japan. This research is also partially supported by Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (A), 14208025, 2003.

References

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control* AC-19, 716–723.
- Anderson, B. D. O., Moore, J. B., 1979. *Optimal Filtering*. Prentice-Hall, New Jersey.
- Harrison, P. J., Stevens, C. F., 1976. Bayesian Forecasting (with discussion). *Journal of Royal Statistical Society* 34, 1–41.
- Harvey, A. C., 1989. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, New York.
- Higuchi, T., Kitagawa, G., January 2000. Knowledge Discovery and Self-Organizing State Space Model. *IEICE TRANS. INF. & SYST.* E83-D (1), 36–43.
- Japan Food Service Association, 2003. Recent business climate of catering industries, <http://www.jfnet.or.jp/data.htm> (in Japanese).
- Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. *Trans. Amer. Soc. Mech. Eng., J. Basic Engineering* 82, 35–45.
- Kitagawa, G., Gersch, W., 1984. A smoothness priors-state space modeling of time series with trend and seasonality. *Journal of the American Statistical Association* 79 (386), 378–89.
- Kitagawa, G., Gersch, W., 1996. *Smoothness Priors Analysis of Time Series*. Springer-Verlag, New York.
- Kitagawa, H., Higuchi, T., Kondo, F. N., 2003. Smoothness prior approach to explore mean structure in large-scale time series. *Theoretical Computer Science* (292), 431–446.
- Kondo, F. N., Kitagawa, G., 2000. Time series analysis of daily scanner sales: extraction of trend, day-of-the-week effect and price promotion effect. *Marketing Intelligence & Planning* 18 (2), 53–66.
- Lilian, G. L., Rangasamy, A., 1997. *Marketing Engineering: Computer-Assisted Marketing Analysis and Planning*. Addison-Wesley Educational Publishers Inc.
- Lilian, G. L., Rangasamy, A., 2002. *Marketing Engineering: Computer-Assisted Marketing Analysis and Planning (2nd Edition)*. Prentice Hall.

- Montgomery, A., 1997. Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data. *Marketing Sci.* 16 (4), 315–337.
- Peppers, D., Rogers, M., 1997. *The One to One Future: Building Relationships One Customer at a Time (One to One)*. Bantam Doubleday Dell Pub.
- Sakamoto, Y., Ishiguro, M., Kitagawa, G., 1986. *Akaike Information Criterion statistics: Mathematics and its application*. D. Reidel, Norwell, Mass.
- West, M., Harrison, J., 1997. *Baysian Forecasting and Dynamic Models*, 2nd Edition. Springer-Verlag, New York.

$y_{1,n}$: Sales of lunches	$y_{2,n}$: Sales of parties	$y_{3,n}$: Sales of the total
$X_{1,n}$: AMAF [#]	$X_{2,n}$: Weather	d_n : Day of the week
	{1, 2, ..., 5}	{1, 2, ..., 7}

Table 1

Time series data sets used in this study. AMAF is an abbreviation of “anticipated mean attendance figure” of events.

	1	0
h_{1n}	National holidays (from Mon to Fri)	the other
h_{2n}	Non-national holidays (from Mon to Thu) preceding a national holiday	the other
h_{3n}	Holidays (Sat, Sun, and national holidays)	the other

Table 2

Conditions for indicator functions of the holiday effect (h_{1n}), the holiday-eve effect (h_{2n}), and the holiday-event augmentation (h_{3n}).

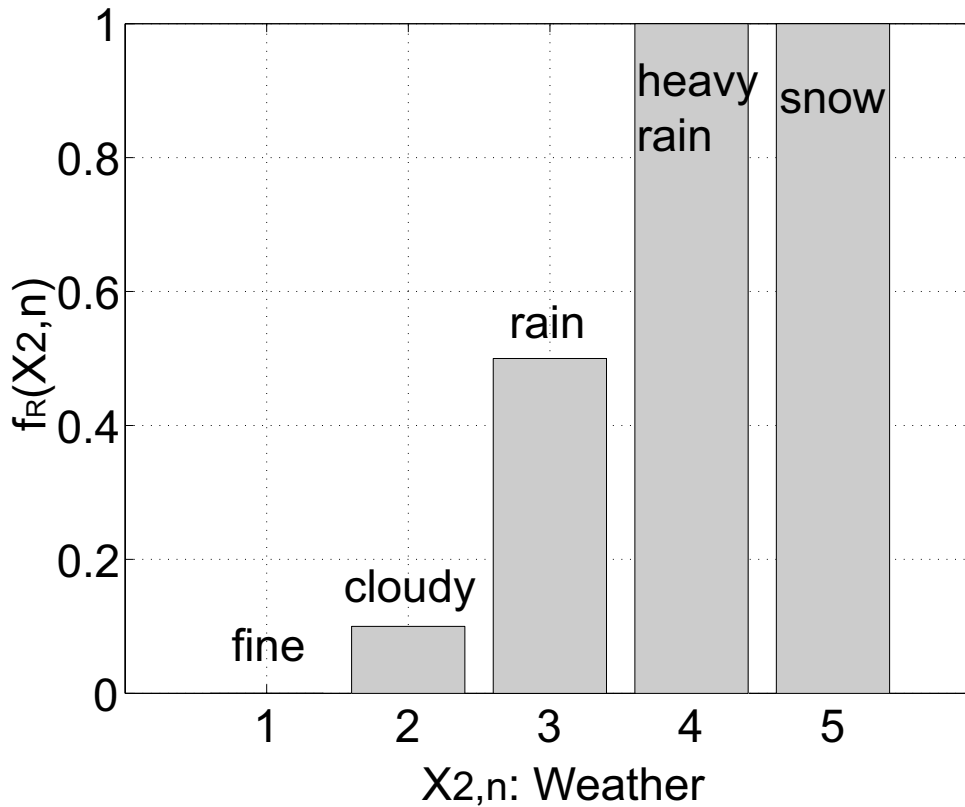


Fig. 1. $X_{2,n}$ vs $f_R(X_{2,n})$

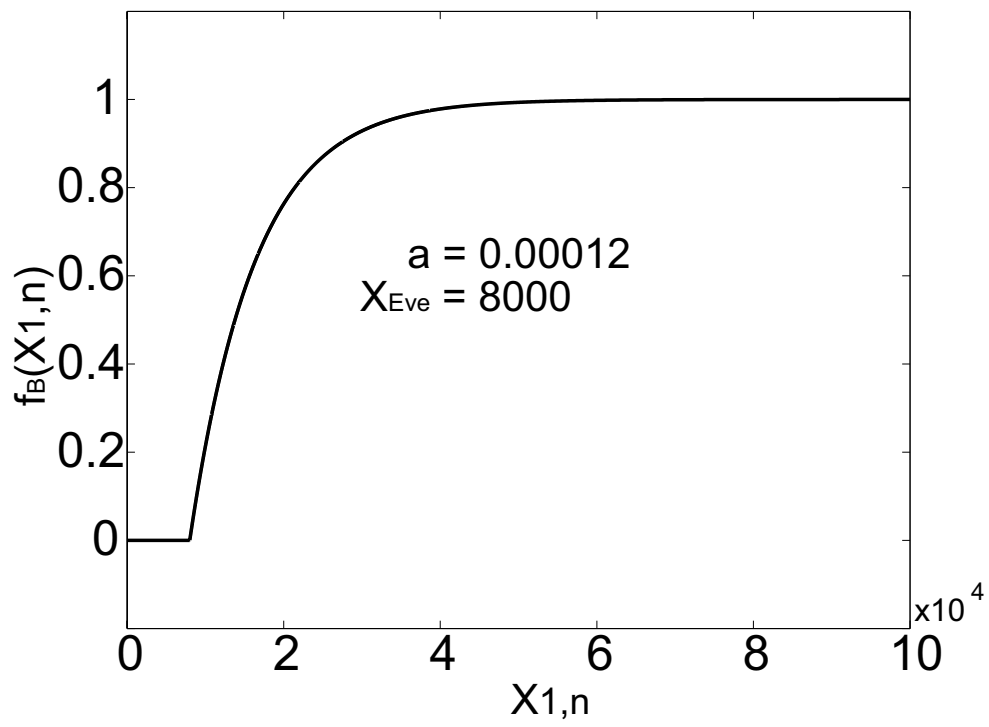


Fig. 2. $X_{1,n}$ vs $f_B(X_{1,n})$

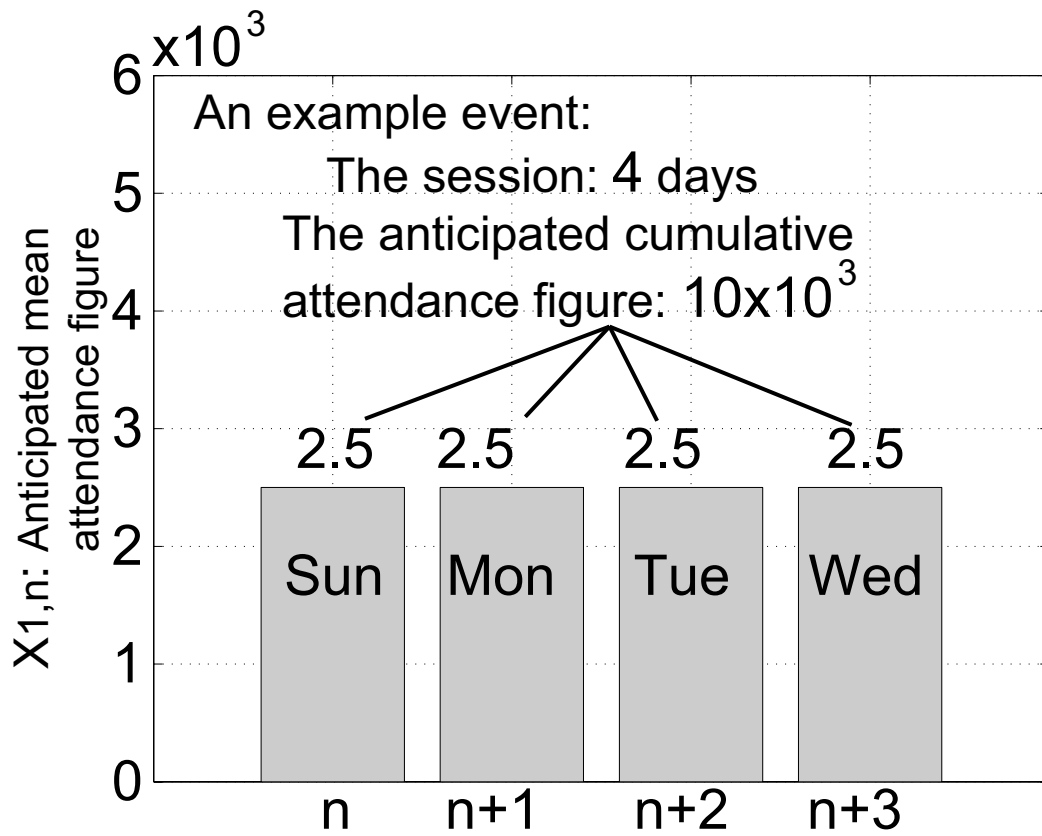


Fig. 3. An example of how to obtain an anticipated mean attendance figure $X_{1,n}$

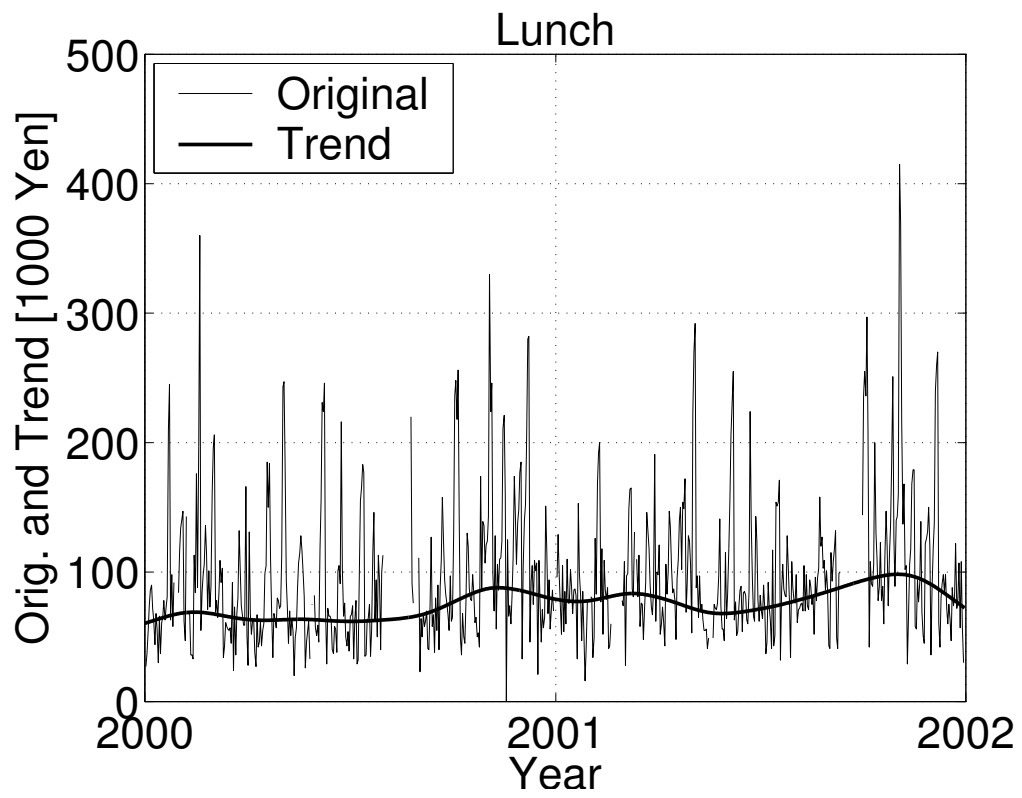


Fig. 4. The original sales of lunches (thin line) and the trend t_n (thick line).

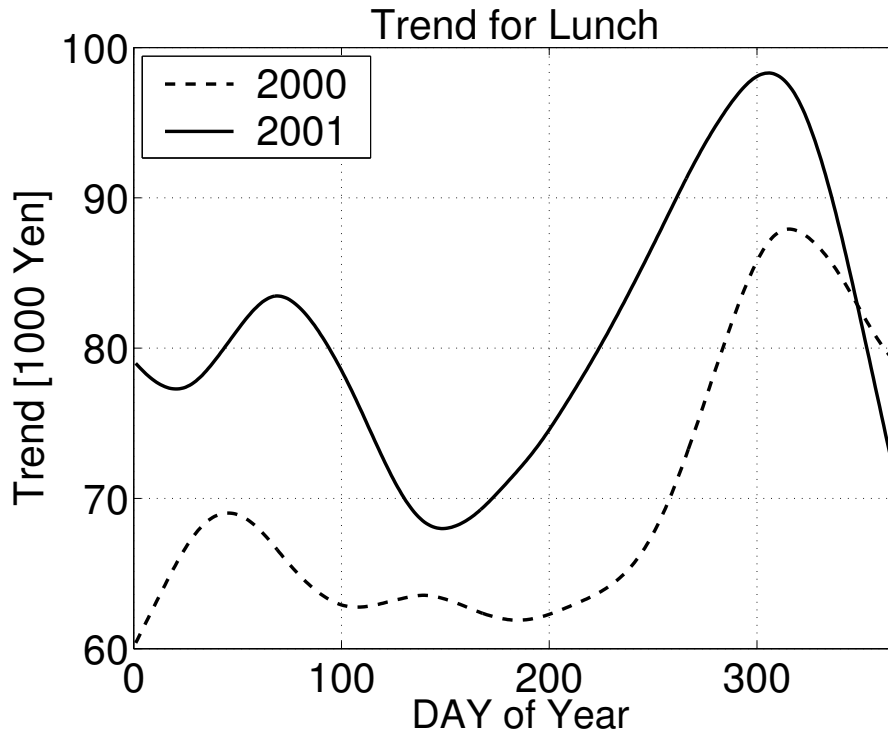


Fig. 5. The year-round variations of the trend during the year 2000 (dashed line) and the year 2001 (solid line).

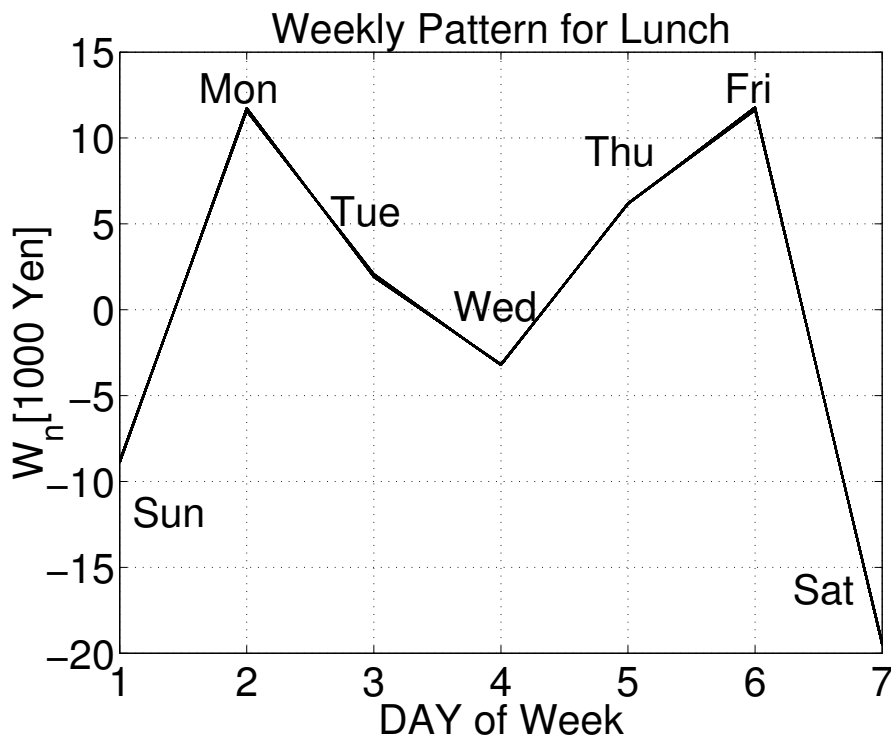


Fig. 6. The day of the week effect w_n for lunch.

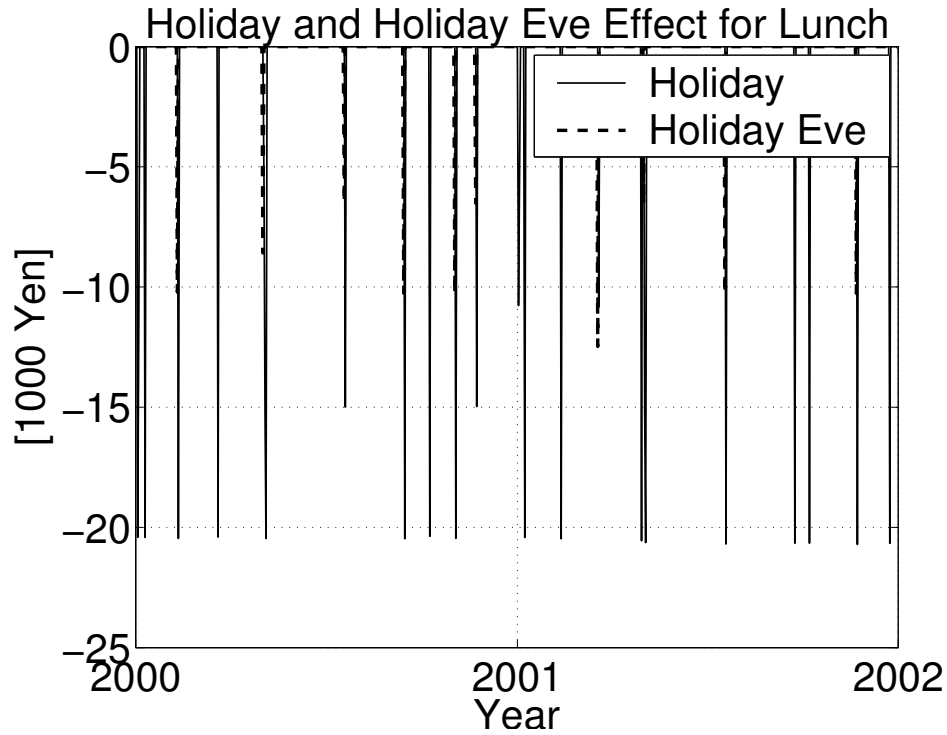


Fig. 7. The holiday effect $h_{1n}\beta_1(w_{Sun,n} - w_n)$ (solid line) and the holiday-eve effect $h_{2n}\{\beta_2(w_{Fri,n} - w_n) + \beta_3(w_{Sat,n} - w_n)\}$ (dashed line) for lunch.

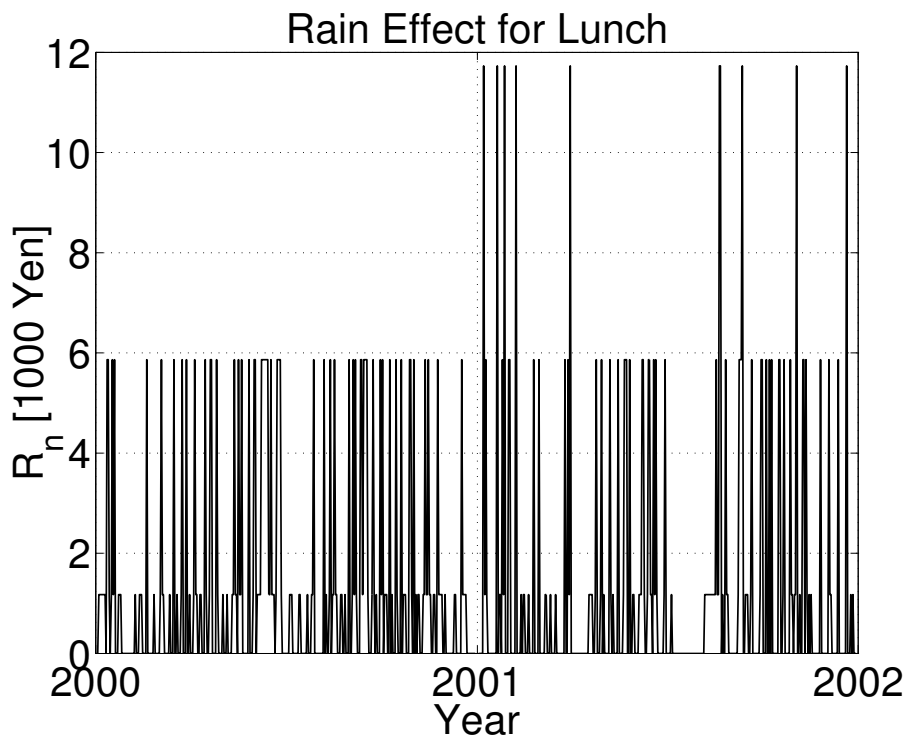


Fig. 8. The rain effect: $R_n = \gamma_R f_R(X_{2,n})$ for lunch.

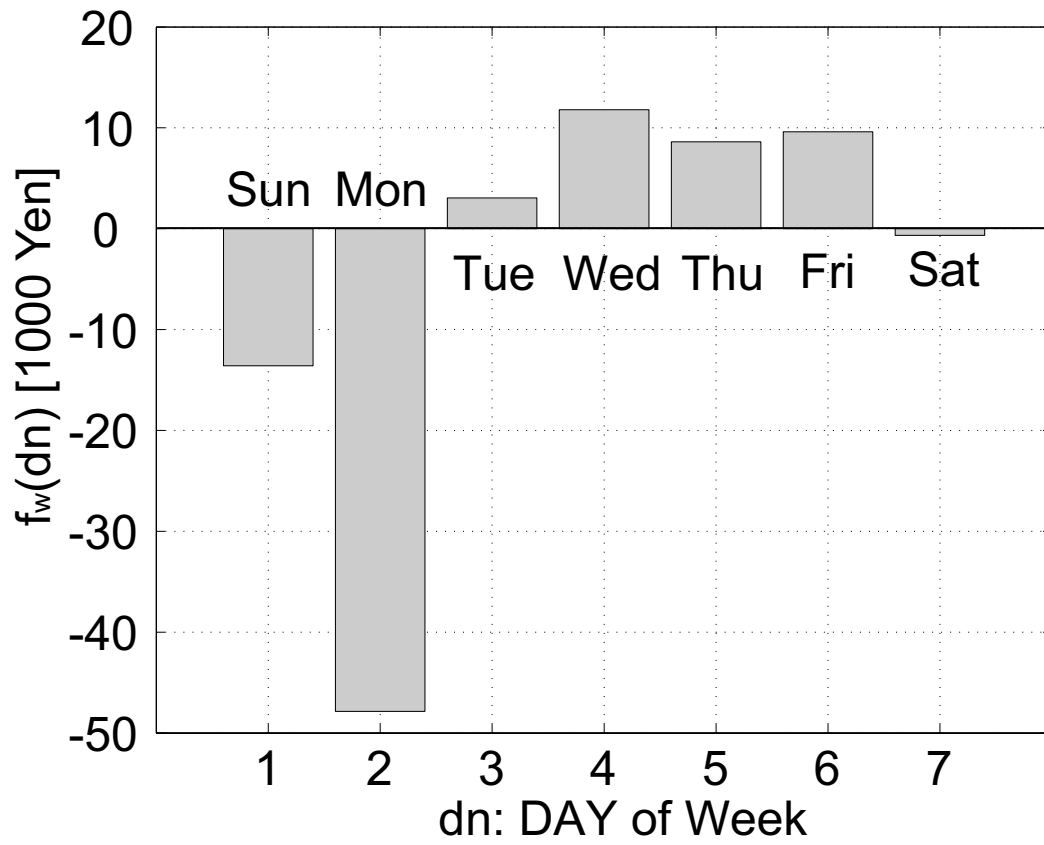


Fig. 9. d_n vs $f_w(d_n)$

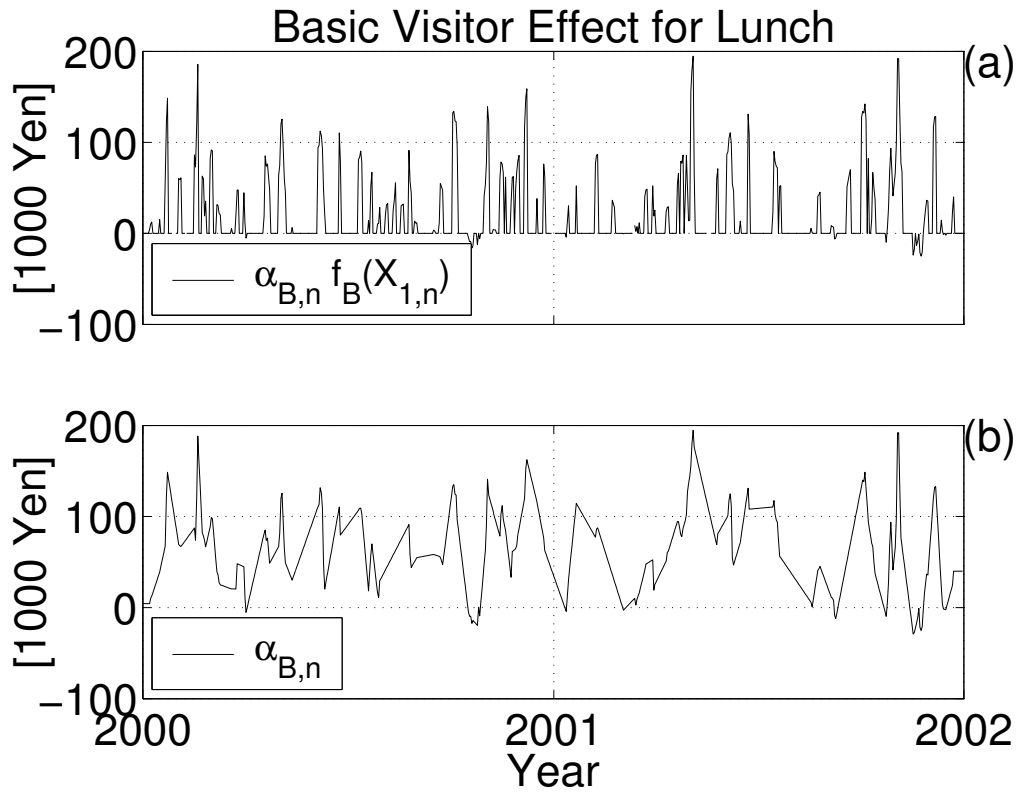


Fig. 10. (a) $\alpha_{B,n} f_B(X_{1,n})$ and (b) $\alpha_{B,n}$

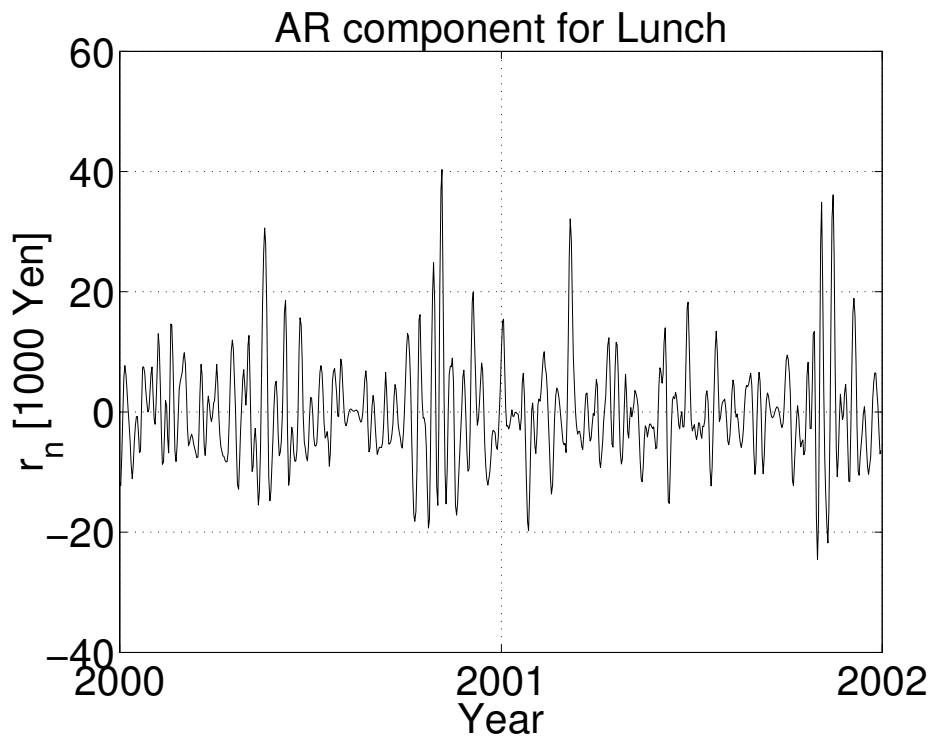


Fig. 11. The stationary second-order AR model component: r_n

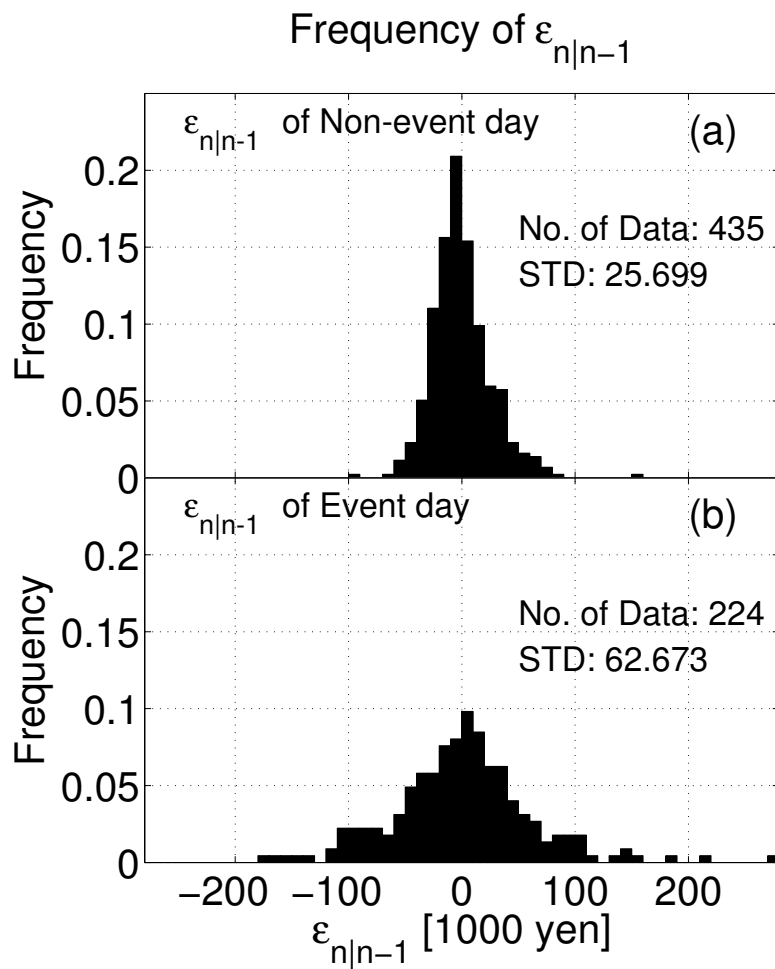


Fig. 12. The frequency of the prediction error ($\varepsilon_{n|n-1}$) for days (a) without event effects and (b) with event effects.

List of MHF Preprint Series, Kyushu University

21st Century COE Program

Development of Dynamic Mathematics with High Functionality

MHF

- 2003-1 Mitsuhiro T. NAKAO, Kouji HASHIMOTO & Yoshitaka WATANABE
A numerical method to verify the invertibility of linear elliptic operators with applications to nonlinear problems
- 2003-2 Masahisa TABATA & Daisuke TAGAMI
Error estimates of finite element methods for nonstationary thermal convection problems with temperature-dependent coefficients
- 2003-3 Tomohiro ANDO, Sadanori KONISHI & Seiya IMOTO
Adaptive learning machines for nonlinear classification and Bayesian information criteria
- 2003-4 Kazuhiro YOKOYAMA
On systems of algebraic equations with parametric exponents
- 2003-5 Masao ISHIKAWA & Masato WAKAYAMA
Applications of Minor Summation Formulas III, Plücker relations, Lattice paths and Pfaffian identities
- 2003-6 Atsushi SUZUKI & Masahisa TABATA
Finite element matrices in congruent subdomains and their effective use for large-scale computations
- 2003-7 Setsuo TANIGUCHI
Stochastic oscillatory integrals - asymptotic and exact expressions for quadratic phase functions -
- 2003-8 Shoki MIYAMOTO & Atsushi YOSHIKAWA
Computable sequences in the Sobolev spaces
- 2003-9 Toru FUJII & Takashi YANAGAWA
Wavelet based estimate for non-linear and non-stationary auto-regressive model
- 2003-10 Atsushi YOSHIKAWA
Maple and wave-front tracking — an experiment
- 2003-11 Masanobu KANEKO
On the local factor of the zeta function of quadratic orders
- 2003-12 Hidefumi KAWASAKI
Conjugate-set game for a nonlinear programming problem

2004-1 Koji YONEMOTO & Takashi YANAGAWA

Estimating the Lyapunov exponent from chaotic time series with dynamic noise

2004-2 Rui YAMAGUCHI, Eiko TSUCHIYA & Tomoyuki HIGUCHI

State space modeling approach to decompose daily sales of a restaurant into time-dependent multi-factors