Retail Strategy to Reduce Waste Keeping High Profit Utilizing Taylor's Law in Point-of-Sales Data

Gen Sakoda, Hideki Takayasu, Misako Takayasu

Abstract—Waste reduction is a fundamental problem for sustainability. Methods for waste reduction with point-of-sales (POS) data are proposed, utilizing the knowledge of a recent econophysics study on a statistical property of POS data. Concretely, the non-stationary time series analysis method based on the Particle Filter is developed, which considers abnormal fluctuation scaling known as Taylor's law. This method is extended for handling incomplete sales data because of stock-outs by introducing maximum likelihood estimation for censored data. The way for optimal stock determination with pricing the cost of waste reduction is also proposed. This study focuses on the examination of the methods for large sales numbers where Taylor's law is obvious. Numerical analysis using aggregated POS data shows the effectiveness of the methods to reduce food waste maintaining a high profit for large sales numbers. Moreover, the way of pricing the cost of waste reduction reveals that a small profit loss realizes substantial waste reduction, especially in the case that the proportionality constant **γ** of Taylor's law is small. Specifically, around 1% profit loss realizes half disposal at **γ**=0.12, which is the actual **γ** value of processed food items used in this research. The methods provide practical and effective solutions for waste reduction keeping a high profit, especially with large sales numbers.

Keywords—Food waste reduction, particle filter, point of sales, sustainable development goals, Taylor's Law, time series analysis.

I. INTRODUCTION

WORLD population is rapidly growing. United Nations predict the population to reach 8.6 billion in 2030 and 9.8 billion in 2050 [1]. Increasing food demand requires the establishment of sustainable food systems. In this context, food waste reduction draws particular interest, since one-third of foods are lost or wasted annually [2]. Thus, Sustainable Development Goals (SDGs) set its target on food waste reduction [3].

Methods for food waste reduction in retail are proposed recently, which considers a statistical property of POS data [4], [5]. Concretely, the sales process is the Poisson process of which fluctuation width is known to be the square root of the mean. However, a recent econophysics study revealed that the fluctuation width follows the mean linearly when the mean is large, which is caused by the fluctuating population [6]. This abnormal fluctuation scaling is known as Taylor's law [7].

G. Sakoda is with the Department of Mathematical and Computing Sciences, School of Computing, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8502, Japan (e-mail: sakoda.g.aa@m.titech.ac.jp).

H. Takayasu is with the Sony Computer Science Laboratories, 3-14-13 Higashi-Gotanda, Shinagawa-ku, Tokyo 141-0022, Japan (e-mail: takayasu@csl.sony.co.jp).

M. Takayasu is with the Institute of Innovative Research, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8502, Japan (e-mail: takayasu.m.aa@m.titech.ac.jp).

Based on this knowledge, a time series analysis method for tracking non-stationary changes of the Poisson parameter considering abnormal fluctuation scaling in sales data is developed, and precise demand estimation is realized [4]. The method is extended for the incomplete observation of demand because of stock-out by integrating the maximum likelihood estimation for censored data [8] into the method. The method for stock determination with pricing the cost of waste reduction is also proposed [5]. Verification with actual POS data in convenience stores proved the effectiveness of the approach. Specifically, the approach decreased the amount of food waste to about a quarter with increasing the profit by 140%, that was confirmed using POS data of foods disposed of frequently about 75% of working days.

The methods are expected to be applicable to sales data in other sectors than retail, such as food manufacturing and supplying industry, and foodservice sectors, since the statistical property of the sales data that follows Taylor's law is assumed to be universal. Meanwhile, the methods were examined with limited POS data in which daily amount of sales are relatively small, typically several tens. Taylor's abnormal fluctuation scaling is more apparent when the sales mean is large. There remains room for examining the methods with large sales numbers, which leads to considering the possibility of the methods for realizing the waste reduction in various sectors.

This study aims to examine the methods for large sales numbers. In section II, the methods and POS data used in this research are described. In section III, the methods are verified using artificial time series and POS data. The final section concludes this study.

II. METHODS

The methods [4], [5] can be divided into three parts namely, the estimation of demand mean value at time t with sales and disposal data, the estimation of demand distribution at time t, and the determination of stock for the next time $t + 1$. The following subsections describe each part of the methods and POS data used in this research.

A. Estimation of the Demand Mean Value

A non-stationary time series analysis method, namely, the Particle Filter [9], [10] is extended to track non-stationary changes of Poisson parameter under the abnormal fluctuation scaling caused by Taylor's law [4].

The Particle Filter is a sort of Monte Carlo simulation which approximates an arbitrary Probability Density Function (PDF) of the state x_t as the distribution of Monte Carlo sample values. The distribution is updated with the observed value y_t in

accordance with the likelihood of each Monte Carlo value. In sales time series, x_t and y_t correspond to demand distribution and observed sales, respectively. Poisson PDF is used as the likelihood function since the sales process is the Poisson process.

The likelihood function is modified considering Taylor's law. The fluctuation term σ of Taylor's law is generally written as:

$$\sigma = \sqrt{\lambda + (\gamma \cdot \lambda)^2} \tag{1}$$

where λ denotes the mean, and γ is the proportionality constant of Taylor's law which is determined with a regression analysis of the mean and the standard deviation of actual data [4].

Poisson PDF is determined without σ, so the Poisson PDF Po is approximated to the Normal PDF N when the mean λ is statistically large in order to incorporate the fluctuation term σ of Taylor's law into the likelihood function.

$$p(y_t|x_t) = \begin{cases} Po(x_t), & if \ x_t < 20 \\ N\left(x_t, \sqrt{x_t + (\gamma \cdot x_t)^2}\right), & otherwise \end{cases} \tag{2}$$

Although the approximation can be done for $x_t \geq 10$ [11], $x_t \geq 20$ is adopted since the skewness of the Poisson distribution at $x_t = 20$ is 30% smaller than that at $x_t = 10$, therefore, closer to 0, that is the value of skewness of the Normal distribution.

The likelihood function is extended to handle the incomplete sales data because of stock-outs. Specifically, the likelihood estimation for censored data is introduced. The following likelihood function for right-censored data takes the expectation of more values than observed into account [8].

$$F(k_t|\theta) = \int_{s_t}^{\infty} f(k|\theta)dk \tag{3}$$

The likelihood function (2) was extended as follows, which considers the cases of censored and non-censored observation.

$$p(y_t|x_t) = \begin{cases} \dfrac{exp\left(\frac{-(y_t-x_t)^2}{2\sigma_t^2}\right)}{\sqrt{2\pi\sigma_t^2}}, & if \ x_t \geq 20 \ and \ non\text{-}censored \\ \dfrac{\int_{m=y_t}^{\infty} exp\left(\frac{-(m-x_t)^2}{2\sigma_t^2}\right)}{\sqrt{2\pi\sigma_t^2}}, & if \ x_t \geq 20 \ and \ censored \\ \dfrac{x_t^{y_t} \cdot exp(-x_t)}{y_t!}, & if \ x_t < 20 \ and \ non\text{-}censored \\ \sum_{m=y_t}^{\infty} \dfrac{x_t^m \cdot exp(-x_t)}{m!}, & otherwise \end{cases} \tag{4}$$

where $\sigma_t = \sqrt{x_t + (\gamma \cdot x_t)^2}$

To track non-stationary changes in time series, the system model of the Particle Filter was also developed. The system model describes a transition of the state from x_t to x_{t+1}, which corresponds to the non-stationary change of demand mean value. Here is the proposed system model:

$$x_t = \delta(x_{t-1} + v_t) \tag{5}$$

$$\delta(x) = \begin{cases} x, & if \ x > 0 \\ 0, & otherwise \end{cases} \tag{6}$$

$$v_t \sim (1-m) \cdot N(0, \alpha \cdot x_{t-1}) + m \cdot U(-\beta, \beta) \tag{7}$$

As shown in (5), the transition from the state x_t to x_{t+1} is stochastically determined with a stochastic variable v_t. The v_t defined in (7) is composed of the Normal distribution $N(0, \alpha \cdot x_{t-1})$ with the mean value zero and the standard deviation $\alpha \cdot x_{t-1}$ $(0 < \alpha)$, and Uniform distribution $U(-\beta, \beta)$ which takes values between $-\beta$ and $\beta (0 < \beta)$. The m characterizes the ratio of superposition $(0 \leq m \leq 1)$. The superposition of two distributions considers that x_t occasionally changes up to β with a small probability m but x_t generally follows the Normal distribution. The range β of the Uniform distribution is selected to be large enough compared to the fluctuation α of the Normal distribution, which allows tracking large non-stationary changes of the state x_t. Delta function (6) is used to limit the value of x_t to be positive since the Poisson parameter needs to be positive. The tilde ($\sim$) in (7) denotes that v_t follows the distribution in the right side of the formula. In this research, m=0.05, α=0.005 and $\beta = 4x_{t-1}$ are adopted by examining Root Mean Squared Error of estimated and assumed parameter value using some artificial time series. The specific procedure of the hyper parameter determination is shown in the Appendix of the reference [4].

The demand mean value λ for each time t is calculated with the median of x_t. The number of Monte Carlo samples is set to N=10,000 and the initial particles are generated by (5)-(7) assuming $x_0 = y_1$, namely sales k_1 for simplicity.

B. Estimation of Demand Distribution

The demand distribution needs to be estimated for the determination of the optimal stock as described in the next subsection. The demand distribution $m(k|\lambda)$ at time t is estimated with the demand mean value λ at time t, and the proportionality constant γ of Taylor's law.

$$m(k|\lambda) = \begin{cases} Po(\lambda), & if \ \lambda < 20 \\ N\left(\lambda, \sqrt{\lambda + (\gamma \cdot \lambda)^2}\right), & otherwise \end{cases} \tag{8}$$

C. Determination of the Optimal Stock

The newsvendor problem formula [12]-[14] tells us the optimal stock s to obtain the maximum profit for the fluctuating demand $(k|\lambda)$. Let the cost c and the price p, the expected profit R(s) is expressed as the following equation.

$$R(s) = p\{\textstyle\sum_{k=0}^{s} km(k|\lambda) + \sum_{k=s+1}^{\infty} sm(k|\lambda)\} - cs \tag{9}$$

Thus, s^* is written with the Inverse of Cumulative Distribution Function (CDF) $M^{-1}(k|\lambda)$,

$$s^* = M^{-1}\left(\frac{c}{p}\Big|\lambda\right) \tag{10}$$

This formula is extended to estimate the dependency of

profit on the ratio of waste reduction, which allows us to determine the stock considering the cost of waste reduction. The expected disposal amount d_t^* for the optimal stock s_t^* is written as.

$$d_t^* = \sum_{k=0}^{s_t^*-1}(s_t^* - k)m(k|\lambda_t) \qquad (11)$$

where $m(k|\lambda_t)$ is the demand distribution shown in (8). Thus, the stock $s(\alpha)_t$ $(0 \leq s(\alpha)_t \leq s_t^*)$ for reducing disposal $\alpha (0 \leq \alpha \leq 1)$ times compared to d_t^* is expressed as:

$$s(\alpha)_t = \operatorname{argmin}(|\alpha d_t^* - \sum_{k=0}^{s-1}(s - k)m(k|\lambda_t)|) \quad (12)$$

The profit $R(s(\alpha)_t)$ is estimated as:

$$R(s(\alpha)_t) = p(s(\alpha)_t - \alpha d_t^*) - cs(\alpha)_t \qquad (13)$$

Thus, the cost of the waste reduction is estimated by examining $R(s(\alpha)_t)/R(s_t^*)$.

The demand distribution $m(k|\lambda_t)$ is assumed to be the Poisson distribution at small λ as shown in (8). Since the Poisson distribution is discrete, the stock $s(\alpha)_t$ obtained with (12) is inevitably discrete, that obstructs fine control of stock. For example, let the demand mean $\lambda = 10$ and the target disposal ratio $0.5 \leq \alpha \leq 1.0$, the stock value $s(\alpha)_t$ can only be 7 or 8. Thus, the Poisson distribution is modified with the Γ function so that the stock can take a real number.

$$m(k|\lambda) = \begin{cases} \dfrac{\lambda^k \cdot exp(-\lambda)}{\Gamma(k+1)}, & if\ \lambda < 20 \\ N\left(\lambda, \sqrt{\lambda + (\gamma \cdot \lambda)^2}\ \right), & otherwise \end{cases} \qquad (14)$$

Although (14) provides a real number of stock, the actual stock value for products at retail needs to be discrete. The probabilistic selection of the two stock values around $s(\alpha)_t$ to approximate the real number of $s(\alpha)_t$ under many trials is proposed.

$$\begin{cases} P(X = \underline{s}) = 1 - (s - \underline{s}) \\ P(X = \underline{s} + 1) = (s - \underline{s}) \end{cases} \qquad (15)$$

where $\underline{s}$ is the integer part of a real number s.

D. POS Data

This research uses POS data of 326 chain stores of a Japanese leading convenience store company, Seven-Eleven Japan Co., Ltd. The POS data covers every purchase at cash registers, and daily records of delivery and disposal for each shop and product, during the 153 days from June to October in 2010. Daily time series on sales, delivery, and disposal for each shop and product are obtained with the POS data.

Sales data of the food products which are wasted more than 75% of sales days are used in this research. There are two reasons for this selection. The first one is that the aim of our research is to reduce food waste and verify that the method can reduce an actual large amount of food waste. The other one is a technical reason. The sales data with many sold-outs do not

represent the underlying demand, which means the estimated demand with the method cannot be verified, especially when the estimated demand is higher than sales. The selected sales data are mainly on food items processed in shops, such as frankfurter sausages, French fries, and fried chickens. The data used in this research are summarized in Table I.

TABLE I
POS DATA USED IN THIS RESEARCH [5]

| Product	Shops	Total Disposal	Total Sales	Price
frankfurter sausages A	11	6030	22,200	150
frankfurter sausages B	8	2670	4291	150
frankfurter sausages C	2	635	1970	150
French fries A	22	8532	21,653	155
French fries B	25	11,939	30,350	155
French fries C	2	921	2615	165
fried chicken A	72	40,140	171,193	165
fried chicken B	33	17,098	152,790	105
fried chicken C	10	4464	14,869	105
grilled chicken A	3	1164	3075	105
grilled chicken B	12	4253	8651	105
grilled chicken C	13	4873	19,272	105
potato croquette	1	452	759	80
skewered beef	1	297	331	120
(Total number)	215	103,468	454,019	-
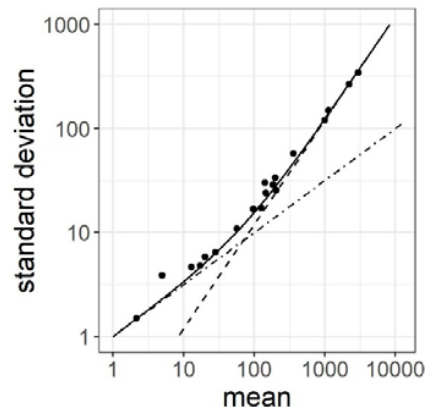
Fig. 1 Mean and standard deviation of the 19 aggregated POS data; Dot-dash line: $\sigma = \sqrt{\lambda}$, Dashed line: $\sigma = \gamma \cdot \lambda$, Solid line: $\sigma = \sqrt{\lambda + (\gamma \cdot \lambda)^2}$, where λ is the mean and γ (=0.12) is the proportionality constant of Taylor's law

In this study, these 215 sales data are aggregated to consider the case that the sales amount is large. Specifically, the sales data is aggregated for each day in various ways, such as for each product (ex. all data of 'frankfurter A'), for each product kind (ex. all data of 'frankfurter'), and all products. Fig. 1 shows the mean and standard deviation of obtained 19 aggregated data, which ensures these aggregated data follows Taylor's law.

This study assumes the cost ratio is 0.7, which is the typical value shown in the company's investor relations [15].
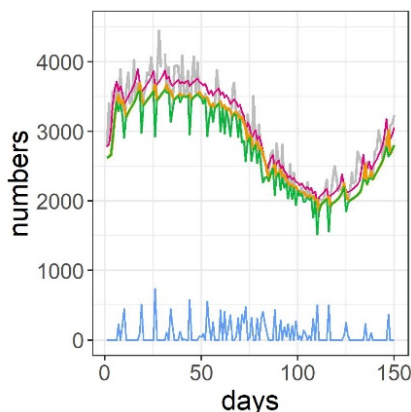
Fig. 2 Sales time series with random number simulation; Green line: sales, Orange line: stock, Blue line: disposal, Gray line: underlying demand, and Magenta line: estimated demand mean

III. RESULTS

A. Verification with Artificial Time Series

Before using actual POS data, artificial time series generated by random number simulation is used to verify the method. Fig. 2 illustrates the simulated sales time series with the method. The gray line indicates underlying demand which is generated by random number simulation assuming the demand distribution (8) and demand mean value in a sine curve shape (the mean value 3000, the amplitude 1800 and the periodicity 150). The proportionality constant γ of Taylor's law is assumed to be 0.1. The orange line shows stock, the green line denotes observed sales, the blue line is disposal, and the magenta line illustrates the estimated mean value of the demand. Here, the stock value for each time t is estimated with (10) using the estimated demand at time t-1. The sales observation is generally censored at the stock value, which can be confirmed in the zero values of disposal. The estimated demand mean value in the magenta line follows the non-stationary changes of demand with censored sales data. The Root Mean Squared Error (RMSE) between the estimated and the assumed mean value is 6.9% on average with 100 sets of random numbers.

Fig. 3 illustrates the dependency between the disposal and the profit which is estimated with the following procedure. The artificial demand time series is generated assuming demand distribution (8). Using (12), the optimal stock for each target disposal ratio between 0.5 and 1.0 is determined. Comparing the demand and the stock values, the amounts of sales and disposal for each time are obtained. The profit is estimated with (13) assuming the cost ratio 0.7 and the price 1. Each series in Fig. 3 is obtained with various proportionality constant γ of Taylor's law. The γ value is 0.12 for the magenta series, 0.05 for the green series, and 0.3 for the orange series. The demand mean value is assumed to be 3000 for these series. The navy series is a reference which demand mean value is 10 where Taylor's law is not obvious. For all series, 10,000 random numbers are used. The dashed line for each series is the theoretical dependency of the profit on the disposal obtained with (12) and (13). The obtained values of the profit and the

disposal fit well to the theoretical dependency.

Fig. 3 provides essential information for waste reduction when the amount of sales is large. Specifically, a small profit loss realizes substantial waste reduction, especially in the case that the proportionality constant γ of Taylor's law is small. While 3.5% profit loss is required to half the disposal compared to that of optimal stock at the sales mean value 10 (the navy line), 1.2% profit loss is needed to half the disposal at the sales mean value 3000 and $\gamma = 0.12$ (the magenta line). The $\gamma=0.12$ is the actual value of processed food items shown in Table I [6].
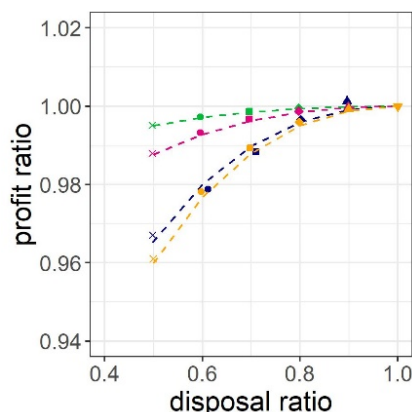

Fig. 3 Simulated dependency of the profit on the disposal for each target disposal ratio between 0.5 to 1.0; Green series: demand mean value λ 3000 and proportionality constant γ of Taylor's law 0.05, Magenta series: λ 3000 and γ 0.12, Orange series: λ 3000 and γ 0.3, and Navy series: λ 10 (Taylor's law is not obvious)
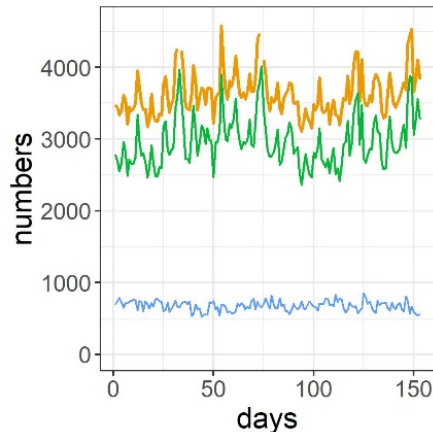

Fig. 4 An example of aggregated POS data; Green line: sales, Orange line: stock, and Blue line: disposal

B. Verification with POS Data

Fig. 4 is an example of aggregated sales time series described in the Methods section. The green line indicates the sales, the orange line shows the stock and the blue line denotes the disposal for each day. In Fig. 4, 103,468 food items in total are disposed of in 153 days. The profit in the period is 9,566,715 yen assuming cost ratio 0.7 and price 150.

Fig. 5 is an example of simulated sales time series with the method using the same aggregated POS data as in Fig. 4. The

gray line indicates the underlying demand. Here, the demand is assumed to be the same as sales in the aggregated POS data. The green line shows the observed sales, the orange line is the stock, the magenta line is the estimated demand mean, and the blue line denotes the disposal. The disposal and the profit in Fig. 5 are 9,449 and 18,157,830 yen, respectively. The disposal is reduced to 9.1% (9,449/103,468) and the profit is increased to 190% (18,157,830/9,566,715) with the method.
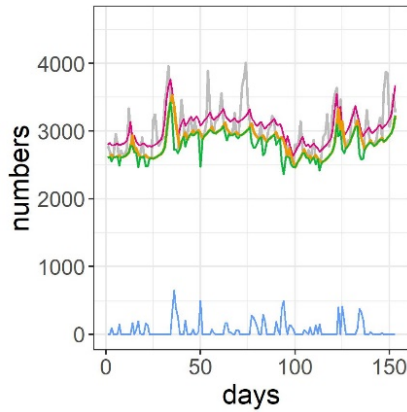

Fig. 5 Simulated sales time series with the method using the aggregated POS data; Green line: sales, Orange line: stock, Blue line: disposal, Gray line: underlying demand, and Magenta line: estimated demand mean
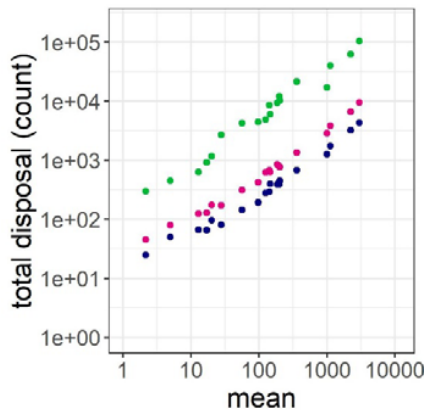

Fig. 6 Disposal of shops and simulated results for the 19 aggregated POS data; Green plot: shop, Magenta plot: the method targeting maximum profit, and Navy plot: the method targeting half disposal

Figs. 6 and 7 show the disposal and the profit of actual shops and simulated results on the 19 aggregated POS data. In Figs. 6 and 7, the green plots are the disposal and the profit of shops, the magenta plots are these of the method targeting maximum profit (i.e., target disposal ratio 1.0), and the navy plots are these of the method targeting half disposal. The target disposal ratio is set to 0.455 to obtain the half disposal to compensate for the deviation between the target and obtained disposal ratio, which is explained in the reference [5]. The method outperforms shops in the disposal and the profit for each 19 aggregated sales time series. Comparing the results of shops

and these of the method targeting maximum profit, the disposal is decreased to 9.5% (median), and the profit is increased to 225% (median). In Fig. 7, some green plots are not shown because the total profit is negative for a large amount of disposal in shops.
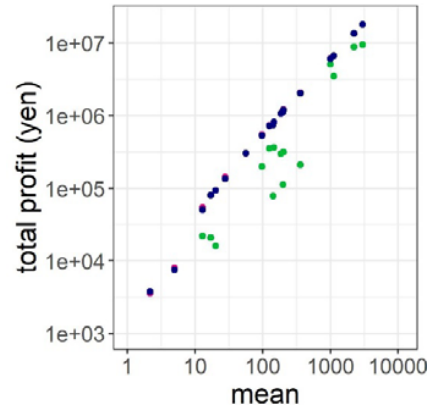

Fig. 7 Profit of shops and simulated results for the 19 aggregated POS data; Green plot: shop, Magenta plot: the method targeting maximum profit, and Navy plot: the method targeting half disposal
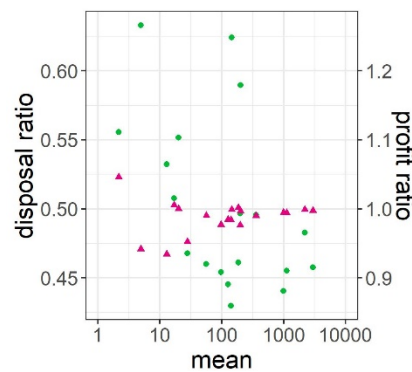

Fig. 8 Profit and disposal of the method targeting half disposal divided by that of the method targeting maximum profit; Green plot: disposal ratio, and Magenta plot: profit ratio

To examine the estimation that a small profit loss realizes substantial waste reduction, the results of the method targeting maximum profit and these of the method targeting half disposal are compared. Fig. 8 shows the ratio of the profit and the disposal which are obtained with dividing the result of the method targeting half disposal by that of the method targeting maximum profit. The typical value (median) of disposal ratio shows 48.3% and that of profit ratio is 99.4%. It is verified that approximately 1% profit loss realizes half disposal at $\gamma = 0.12$.

IV. CONCLUSION

The non-stationary demand estimation and stock determination methods are examined in the case of large sales numbers, where the abnormal fluctuation scaling caused by Taylor's law is apparent. Artificial time series and actual POS data verified the effectiveness of the methods for food waste reduction with maintaining a high profit. The way of pricing the

cost of waste reduction shows that a small profit loss realizes substantial waste reduction, especially in the case that the proportionality constant γ of Taylor's law is small. POS data verified that around 1% profit loss realizes half disposal at γ=0.12 which is the actual γ value of processed food items used in this research. The method is expected to reduce waste keeping a high profit in retail and the other sectors especially with large sales numbers.

REFERENCES

[1] United Nations, "World Population Prospects the 2017 Revision," USA: United Nations, 2017.
[2] Food and Agriculture Organization of the United Nations, "Food Wastage Footprint—Impacts on Natural Resources," Rome: Food and Agriculture Organization of the United Nations, 2013.
[3] United Nations, "Transforming Our World: The 2030 Agenda for Sustainable Development," New York: United Nations, 2015.
[4] G. Sakoda, H. Takayasu, and M. Takayasu, "Tracking Poisson Parameter for Non-Stationary Discontinuous Time Series with Taylors Abnormal Fluctuation Scaling," Stats, vol. 2, no. 1, pp. 5569, Jan. 2019.
[5] G. Sakoda, H. Takayasu, and M. Takayasu, "Data Science Solutions for Retail Strategy to Reduce Waste Keeping High Profit," Sustainability, vol. 11, no. 13, p. 3589, Jun. 2019.
[6] G. Fukunaga, H. Takayasu, and M. Takayasu, "Property of fluctuations of sales quantities by product category in convenience stores," PLoS One, vol. 11, no. 6, pp. 119, 2016. PLoS ONE 2016, 11, e0157653.
[7] L. R. Taylor, "Aggregation, variance and the mean," Nature, vol. 189, pp.732–735, 1961.
[8] E. T. Lee, J.W.Wang, "Statistical Methods for Survival Data Analysis," USA: WILEY, 2013, pp. 133-205.
[9] G. Kitagawa, "Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models," J. Comput. Graph. Stat, vol.5, pp.1–25, 1996.
[10] N. J. Gordon, D.J. Salmond, A.F.M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," IEE Proc.F, vol. 140, pp.107–113, 1993.
[11] S. R. Cherry, J.A. Sorenson, M.E. Phelps, "Physics in Nuclear Medicine," 4th ed. Amsterdam: Elsevier, pp. 126–128, 2012.
[12] E. L. Porteus, "Stochastic Inventory Theory. Handbooks in Operations Research and Management Science," Amsterdam: Elsevier, vol. 2, pp. 605–652, 1990.
[13] M. Khouja, "The single-period (news-vendor) problem: Literature review and suggestions for future research," Omega, vol. 27, pp. 537–553, 1999.
[14] Y. Qin, R. Wang, A.J. Vakharia, Y. Chen, M.M.H. Seref, "The newsvendor problem: Review and directions for future research," Eur. J. Oper. Res., vol. 213, pp. 361–374, 2011.
[15] Seven & i Holdings Co., Ltd., "Corporate Outline 2011," Available online:https://www.7andi.com/library/dbps_data/_template_/_res/en/ir/library/co/pdf/2011_07.pdf (accessed on 8 September 2019).