International Journal of Networked and Distributed Computing, Vol. 5, No. 4 (October 2017) 233–242

Improvement and Evaluation of Estimation of Time Series Data of Daily Life

Teruhisa Hochin
Information and Human Sciences, Kyoto Institute of Technology, Goshokaido-cho, Matsugasaki, Sakyo-ku
Kyoto 606-8585, Japan
E-mail: hochin@kit.ac.jp

Hiroki Nomiya
Information and Human Sciences, Kyoto Institute of Technology, Goshokaido-cho, Matsugasaki, Sakyo-ku
Kyoto 606-8585, Japan
E-mail: nomiya@kit.ac.jp

Abstract

This paper improves the estimation of the amounts of sewage flow, which is one of daily life data, in order to manage them efficiently. The amounts of flow of a typical day are tried to be adjusted to those of a non-regular day. A typical (non-regular, respectively) day is a non-rainy day having good data and no (a few) outliers. The values for the adjustment are tried to be estimated by using the multiple regression analysis. It is shown that the estimation can be improved, and these values can be estimated by using the temperature of that day, the amount of the rain fall of the previous day, and the day type, which distinguishes a weekday, Saturday, Sunday, and a national holiday. The estimation is tried to be used in estimating the data of a regular day. It is experimentally shown that the estimation works well.

Keywords: Estimation, Daily life data, Multiple regression analysis, Day type, Rainfall, Temperature.

1. Introduction

There are many time series data of daily life. Traffic volume, electricity usage, and air temperature change are such data. Sewage flow is one of those data. Sewage flow must be considered to in the sewage treatment.

Sewage treatment is the process of removing contaminants from wastewater.[1] Basic sewer systems were used even in ancient Mesopotamia, the Indus Valley civilization, Ancient Crete and Greece, and the Roman Empire.[1] Sewage treatment is inevitable for modern society, and is indispensable for the healthy life of people.

Energy requirement for sewage management is very high. A lot of energy is consumed. As efficient control of sewage management results in saving energy, it is considered to be very important. It is said that clarifiers and mechanized secondary treatment are more efficient under uniform flow conditions.[1] For making the flow uniform, the estimation of the load of sewage flow is required. If the load of sewage flow can be known in advance, the plan of driving the plant can be fit to the optimum one. The load of sewage flow is based on the amounts and the quality of sewage flow. They are strongly required to be estimated.

The amounts of sewage flow were tried to be estimated in order to manage them efficiently.[2] Here, only the days which are not rainy, and have the data good enough to process were treated. The days having good data are called normal days. Normal days are divided into typical days and irregular ones. The latter have outliers, and the former do not. Typical days are again classified into regular days and non-regular ones based on the existence of outliers. The amounts of flow on regular days are almost those of the typical day. The amounts of flow of a typical day were tried to be adjusted to those of a non-regular day. The values for the adjustment were also tried to be estimated through the multiple regression analysis. Although it is shown that these values could be estimated by using the temperature and the amount of rain fall of the previous day, the estimation accuracy was not good.

This paper improves the estimation accuracy of the values used in adjusting the amounts of flow of a typical day to that of a non-regular day. It is shown that the estimation accuracy can be improved by using the values other than the previous ones as explanatory variables. Although these explanatory variables are different from the previous ones, these are still of the temperature of that day and the amount of rain fall of the previous day. This paper also evaluates the degree of the improvement. First, the estimation accuracy of the regression formula is evaluated. It is shown that the estimation accuracy is improved, and the estimation errors become small. Next, we try to adapt the estimation of the values of a non-regular day from those of the typical one to those of a regular day. The adaptation is evaluated by using another data set. It is shown that the estimation works well.

The structure of the remainder of the paper is as follows: Section 2 describes sewage management. Section 3 summarizes the categorization and the estimation result of sewage flow at the previous research.[2] Section 4 improves the estimation of the sewage flow. Section 5 evaluates the estimation revised. Section 6 gives some considerations. Finally, Section 7 concludes the paper.

2. Sewage Treatment

2.1. *Overview*

Sewage treatment is the process of removing contaminants from wastewater, primarily from household sewage.[1] Physical, chemical, and biological processes are included in order to remove contaminants and to produce treated wastewater which is environmentally safe. Semi-solid waste or slurry is usually produced. It is called sewage sludge. It must be treated in order to be suitable for disposal or land application.

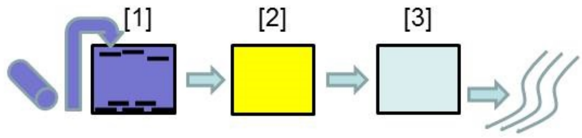


Fig. 1. Process steps of sewage treatment.

Sewage treatment may also be called wastewater treatment. It, however, is a general term. It can also be applied to purely industrial wastewater. The sewer system also carries industrial effluent to the sewage treatment plant for most cities. It has usually received pretreatment at the factories in order to reduce the pollutant load. If the sewer is a combined one, then it will also carry urban runoff including storm water to the sewage treatment plant.

2.2. *Process Steps*

The end of the sewage treatment is to produce an effluent that will not harm the surrounding environment when it is discharged to the environment. It is strongly required to prevent pollution from being released into the surrounding environment, and to contaminate it.

Sewage treatment generally has three stages: primary, secondary, and tertiary treatments as shown in Fig. 1.

(1) Primary treatment

In the primary treatment, the sewage is held in a quiescent basin for a while. Heavy solids sink to the bottom of the basin, while oil, grease, and lighter solids float to the surface. The sunk and floating materials are removed. The remaining liquid may be discharged. It may also be subjected to the secondary treatment.

During very heavy rainfall event, the secondary and tertiary treatment systems should be bypassed in order to protect them from hydraulic overloading, and the mixture of sewage and storm water. For this purpose, some sewage treatment plants have a bypass after the primary treatment unit. Storm water is treated only in the primary treatment.

(2) Secondary treatment

Dissolved and suspended biological matter is removed in the secondary treatment. Indigenous, water-borne micro-organisms in a managed habitat are typically used. Removing the micro-organisms from the treated water is required before the discharge or the process of the tertiary treatment.

(3) Tertiary treatment
In the tertiary treatment, the water processed through the primary and the secondary treatments is processed in order to be disposed into a highly sensitive or fragile ecosystem, which includes estuaries, low-flow rivers, coral reefs, and so on.

The water treated is chemically or physically disinfected before discharged into a stream, river, bay, lagoon, or wetland. It may also be used for the irrigation of a golf course, green way, or park. It can be used for groundwater recharge or agricultural purposes if it is sufficiently clean.

2.3. *Energy Requirement*

Around thirty percent of the annual operating costs are usually required for energy for the conventional sewage treatment plant.[1] The energy requirements depend on the type of treatment process as well as wastewater load. The wastewater load is based on the amounts and the quality of sewage flow.

In the conventional secondary treatment process, most of the electricity is used for aeration, pumping systems, and the equipment for the dewatering and drying of sewage sludge. Advanced wastewater treatment plants require more energy than the plants only achieving primary and secondary treatments.

It is said that clarifiers and mechanized secondary treatment are more efficient under uniform flow conditions.[1] In order to make the flow uniform, the estimation of the wastewater load is required. If the load of sewage flow can be known in advance, the plan of driving the plant can be fit to the optimum one.

3. Estimation of Sewage Flow

The sewage flow was firstly categorized into several types.[2] The typical sewage flow was then clarified. The sewage flow other than the typical one was tried to be estimated.[2] This section describes the method and the results obtained in the previous research.[2]

3.1. *Data Used*

The sewage flow data collected in Oct. 1st, 2015 to Oct. 31st, 2016 were used.[2] The average amount of flow is hourly measured. There are twenty-four data points a day. The day having a missing value is excluded from the day treated here.

3.2. *Categorization of Sewage Flow*

3.2.1. *Normal Days*

Only non-rainy days are treated in order to make the analysis easy. Here, the day having the rainfall amount less than 5 mm per hour is considered as a non-rainy day.

The amount of flow may suddenly become very large. The day including such a sudden huge amount of flow is not treated in this paper. As the average amount of flow is around 2,000 as described later, the day including the flow amount of 3,000 or more is not treated. These days are called normal days. We obtained 186 normal days. The amounts of flow are averaged. These averaged amounts of flow are called the tentative typical amounts of flow.

3.2.2. *Typical and Irregular Days*

Normal days are divided into two categories: typical days and irregular ones. Irregular days are the days including outliers from the normal days. The other days are typical days. For selecting typical days, we adopted three criteria: (1) A typical day has less than 20,000 of the sum of the positive differences of flow amounts, (2) It has less than 2,000 of the sum of the absolute values of their negative differences, and (3) Correlation to the tentative typical amounts of flow is more than 70 percent. 127 typical days were obtained. The amounts of flow of these days are averaged to obtain the typical amounts of flow. The typical amount of flow is shown in Fig. 2. The amount of flow changes according to our daily life. That is, the amount is small in early morning, and has peaks around 9AM and 10PM.

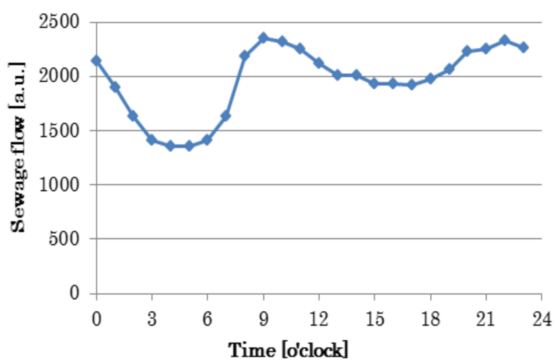


Fig. 2. Typical flow pattern not considering day types.

3.2.3. *Regular and Non-Regular Days*

Moreover, the typical days are divided into regular days and non-regular days. For selecting regular days, we also adopted three criteria. The first two criteria are the same as those of (1) and (2) adopted in selecting typical days. The last one is the criterion that correlation to the typical amounts of flow is more than 70 percent. It is slightly different from that of (3) in selecting typical days. 94 days are selected as regular days. The other typical days are non-regular days. 33 days are selected as non-regular days.

3.3. *Day Type*

Until here, the type of day is not considered. Our daily life differs depending on the weekday or the weekend. Here, days are categorized into the weekday, Saturday, Sunday, and the national holiday. These are called types of days. Typical, regular, and non-regular days are decided in considering the type of day. These are 130, 101, and 29 days, respectively.

The typical amounts of flow in considering the type of day are shown in Fig. 3. The shape of flow is quite similar to that shown in Fig. 2. There are, however, some differences. The amount begins to increase at six o'clock on a weekday, while it does at seven on days of a week end and a national holiday. A peak is at nine o'clock on weekdays, while it is at ten o'clock on the other days. The lowest amount on a national holyday is around 1,200, while it is around 1,400 on the other days. The average amount of flow on weekdays is larger than that on the other days. The amount of the night peak on weekdays is larger than that on the other days. Although the peak at the night begins to decrease at 23 o'clock on weekdays, it remains even at midnight on the other days.

3.4. *Estimating Sewage Flow*

3.4.1. *Adjusting Non-Regular Data*

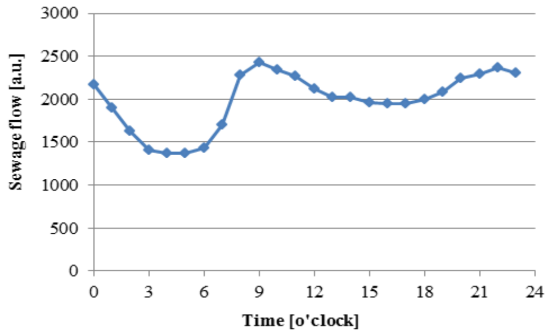
As the amounts of flow of the regular days are almost the same as those of the typical day, the amounts of flow of the typical day can be used as those of the regular days.

As it is not true for those of the non-regular days, those are tried to be obtained by adjusting the amounts of flow of the typical day to those of a non-regular one. Here, we consider the scaling and the offset. The transformation of the amount of flow of the typical day *d(t)* to that of a non-regular day *d'(t)* at the time *t* is represented by Eq. (1).

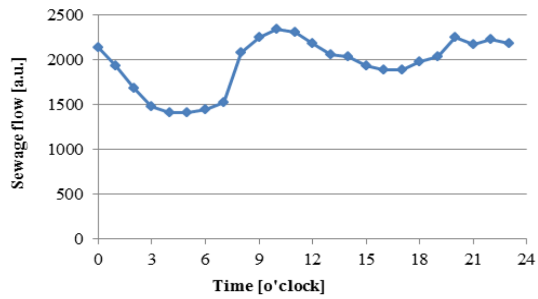
$$d'(t) = m * d(t) + ofst \qquad (1)$$

Here, *m* is for scaling, and *ofst* is an offset value.

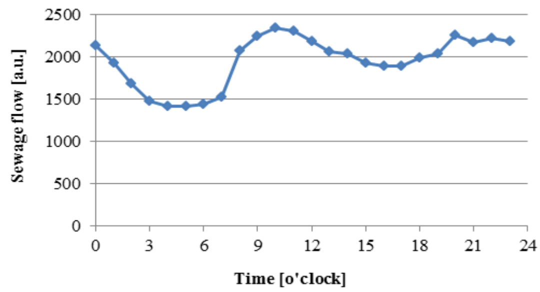
By using the least-squares method, we could decide



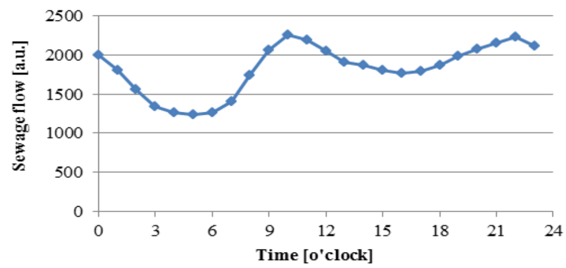
(a) Weekday



(b) Saturday



(c) Sunday



(d) National Holiday

Fig. 3. Typical flow pattern considering day types.

Table 1. Explanatory Variable

Type	*m*	*ofst*
Considering no day type	Amount of rain fall of the previous day Eighth power of the mean temperature Day type code	Highest temperature Square root of the amount of rain fall of the previous day Day type code
Considering day types	Third power of the mean temperature	Square root of the amount of rain fall of the previous day Second power of the highest temperature

the values of *m* and *ofst* for each day.

For the days not considering the type of day, the mean correlation of the amounts of the original non-regular days and those of the adjusted ones is 0.95.

For the days considering the type of day, the mean correlation of the amounts of the original non-regular days and those of the adjusted ones is 0.96.

3.4.2. *Estimating Non-Regular Data*

Although the values of *m* and *ofst* could precisely be estimated, these values can be calculated only after the amounts of flow of a day are known. As the aim of this study is to estimate the amounts of flow on a future day, it is not enough to only obtain the values of *m* and *ofst*.

The weather data including weather, temperatures, humidity, and rainfall amount are additionally used. These are used as explanatory variables of the regression analysis.

(1) Estimation of Non-Regular Data without Considering the Day Type

The value *m* can be estimated by using the amount of rain fall of the previous day, the eighth power of the mean temperature, and the day type code as explanatory variables. Here, the day type code is the number that Monday to Sunday is represented with one to seven, respectively, and a national holiday is represented with eight. The explanatory variables used are listed in Table 1. The multiple correlation coefficient of the estimation was 0.56.

The value of *ofst* could be estimated by using the highest temperature, the square root of the amount of rain fall of the previous day, and the day type code. The multiple correlation coefficient of the estimation was 0.49.

(2) Estimation of Non-Regular Data Considering the Day Type

The value *m* can be estimated by using the third power of the mean temperature as an explanatory variable. The multiple correlation coefficient of the estimation was 0.58.

The value of *ofst* could be estimated by using the square root of the amount of rain fall of the previous day, and the second power of the highest temperature. The multiple correlation coefficient of the estimation was 0.66.

3.5. *Knowledge Obtained*

Although the estimation accuracy of the values of *m* and *ofst* were not high, those in considering the day type are slightly better than those not in considering it. Considering the day type seems to be effective for the precise estimation of the amounts of flow. In the case that the day type is not considered, the day type code is used as an explanatory variable in the estimation. This fact supports that distinguishing the day type is required in estimating the amounts of flow.

For the estimation of the value *m*, the mean temperature is commonly used as an explanatory variable with and without considering the day type. It is considered that the mean temperature may affect the scaling of the amounts of flow.

For the estimation of the value *ofst*, the amount of rain fall of the previous day and the highest temperature are commonly used as explanatory variables with and without considering the day type. These two factors may affect the offset of the amounts of flow.

4. Improvement of Sewage Flow Estimation

In order to improve the estimation accuracy, we tried to use other data as explanatory variables.

4.1. *Estimation of Non-Regular Data*

4.1.1. *Estimating Non-Regular Data without Considering the Day Type*

The value *m* can be estimated better by using the square root of the amount of the rain fall of the previous day, the square root of the mean temperature, and the day type code as explanatory variables. The explanatory variables adopted here are listed in Table 2. The multiple correlation coefficient of the estimation was improved to 0.62.

The value of *ofst* can be estimated better by using the same values as those used in estimating the value of *m* as explanatory variables. The multiple correlation coefficient of the estimation was improved to 0.64.

Table 2. Revised Explanatory Variable

Type	*m*	*ofst*
Considering no day type	Square root of the amount of rain fall of the previous day Square root of the mean temperature Day type code	Square root of the amount of rain fall of the previous day Square root of the mean temperature Day type code
Considering day types	Amount of rain fall of the previous day Second power of the highest temperature	Amount of rain fall of the previous day Highest temperature

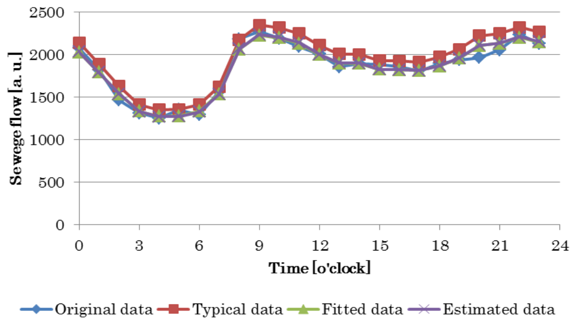
Examples of estimation without considering the day type are shown in Fig. 4. Fig. 4 (a), (b), and (c) show the estimation results having the minimum estimation error, the error around the average one, and the maximum one, respectively. The estimated data as well as the original, the typical, and the adjusted data are shown. It can be seen in Fig. 4 that the original data are a little bit apart from the typical ones, while the adjusted and the estimated ones are very close to the original one. Even in the case of the maximum error shown in Fig. 4 (c), the adjusted and the estimated data are very close to the original data. The amounts of flow seem to be precisely adjusted and estimated.

4.1.2. *Estimating Non-Regular Data Considering the Day Type*

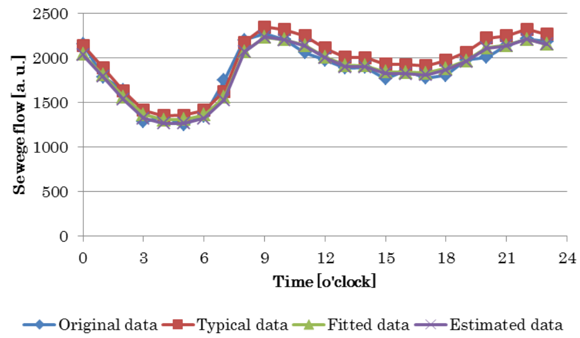
The value *m* can be estimated better by using the amount of the rain fall of the previous day and the second power of the highest temperature as explanatory variables. The multiple correlation coefficient of the estimation was improved to 0.62.

The value of *ofst* can be estimated by using the amount of the rain fall of the previous day and the highest temperature as explanatory variables. The multiple correlation coefficient of the estimation was 0.60.

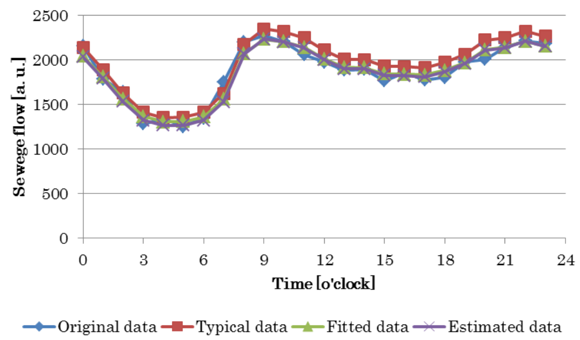
Examples of estimation considering the day type are shown in Fig. 5. Fig. 5 (a), (b), and (c) show the estimation results having the minimum estimation error, the error around the average one, and the maximum one, respectively. In Fig. 5 (a) and (b), the typical and the adjusted data are very similar to the original data. In Fig. 5 (c), the time of the first peak of the amount of flow is shifted by comparing with that of the original one. Adjustment of the amount of flow seems to try to fit the amount of flow to the average one. Shifting the peak is not considered in the adjustment as shown in Eq. (1). This shift should be considered in future.



(a) Data with the minimum adjustment error

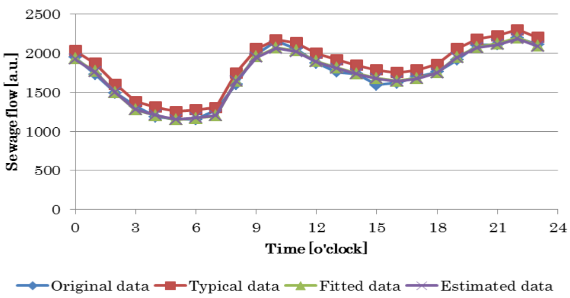


(b) Data with the error around the average one

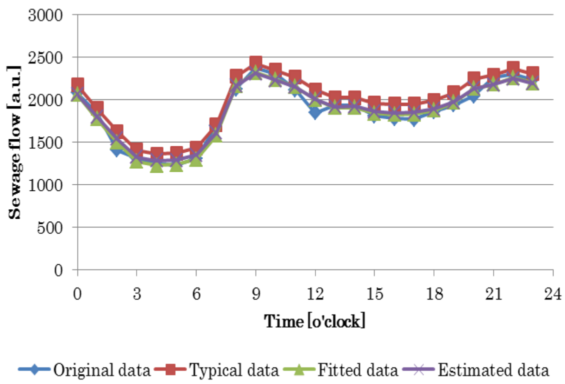


(c) Data with the maximum adjustment error

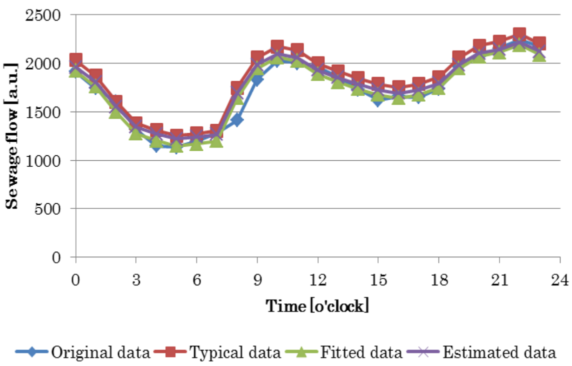
Fig. 4. Estimation of flow patterns without considering day types.



(a) Data with the minimum adjustment error



(b) Data with the error around the average one



(c) Data with the maximum adjustment error

Fig. 5. Estimation of flow patterns considering day types.

4.2. *Considerations*

For the estimations of the values *m* and *ofst*, the amount of rain fall of the previous day and the temperature are commonly used as in the previous estimations. It is considered that these values are effective in estimating the amount of sewage flow.

Considering the day type also seems to be effective for the precise estimation of the amounts of flow. In the case that the day type is not considered, the day type code is used as an explanatory variable in the revised estimation. This fact supports that distinguishing the day type is required in estimating the amounts of flow.

In considering no day type, the explanatory variables used in estimating the value *m* are surprisingly the same as those used in estimating the value *ofst*.

5. Evaluation

5.1. *Evaluation Method*

The estimation revised is evaluated in order to show the degree of the improvement. Adaptability of the estimation revised is also examined.

(1) Evaluation 1
The degree of the improvement of the estimation from the fitted values is evaluated. The multiple correlation coefficients of the estimation and the relative errors between the fitted values and the estimation ones are used in evaluating the improvement degree.

(2) Evaluation 2
The adaptability of the estimation is examined by using another data set. This data set includes the data of the sewage flow collected in November 1st, 2016 to April 23rd, 2017.

The values of sewage flow are estimated by using the values of the typical data, and the estimation formulas obtained from the previous data set, even though these formulas are for estimating non-regular data from those of the typical data. The values are estimated both in considering no day type and in considering day types.

The correlation values of the original values and the estimated ones, and their relative error rates are used in evaluating the adaptability.

5.2. *Evaluation Result*

5.2.1. *Evaluation 1*

The multiple correlation coefficients of the previous and the improved estimations are shown in Table 3. Both of the values of the multiple correlation coefficients of the estimation of *m* and that of *ofst* in considering no day type increased. The value of the multiple correlation coefficient of the estimation of *m* increases. Although the value of the multiple correlation coefficient of the

Table 3. Multiple Correlation Coefficients

Type	Considering no day type		Considering day types	
	m	*ofst*	*m*	*ofst*
Original	0.56	0.49	0.58	0.66
Revised	0.62	0.64	0.62	0.60

Table 4. Relative Error Rate

Estimation	Considering no day type	Considering day types
Original	8.8 %	32.8 %
Revised	5.3 %	6.3 %

Table 5. Correlation and Relative Error Rate

Type	Correlation	Relative error rate
Considering no day type	0.91	48 %
Considering day types	0.93	44 %

estimation of *ofst* decreases, it is still equal to 6.0.

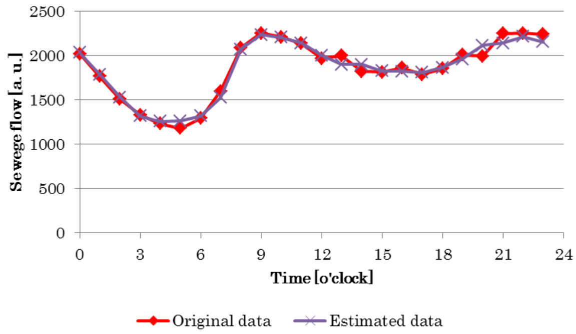
The mean values of the relative error rates between the fitted values and the estimation ones are shown in Table 4. Both of the relative error rate in considering no day type and that in considering day types became small. Especially, the relative error rate in considering day types became drastically small.

5.2.2. *Evaluation 2*

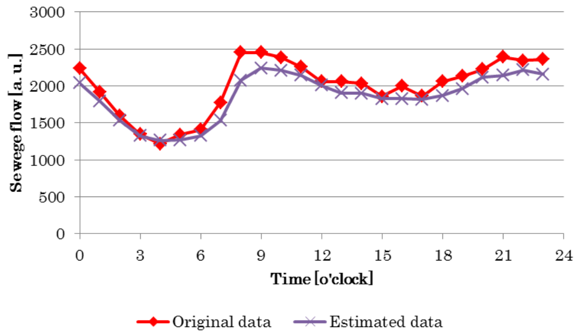
There are 174 days in the data set used in the evaluation. 148 days are non-rainy ones. We could find 73 normal days in them. These were categorized into 27 typical days and 46 irregular ones. Typical days were categorized into 26 regular days and only one non-regular day.

The correlation values of the original values and the estimated ones, and their relative error rates are shown in Table 5. Both of the correlation value and the relative error rate in considering day types are better than those in considering no day type.

Results of estimation without considering the day type are shown in Fig. 6. Fig. 6 (a), and (b) show the estimation results having the minimum estimation error, and the error around the average one, respectively. The values of flow are estimated very well in Fig. 6 (a). In Fig. 6 (b), the estimated values are a little bit apart from the original ones.

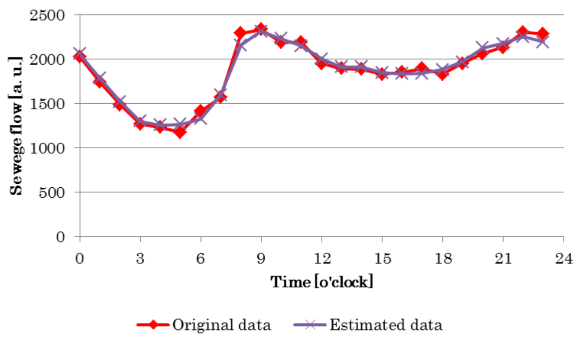


(a) Data with the minimum estimation error

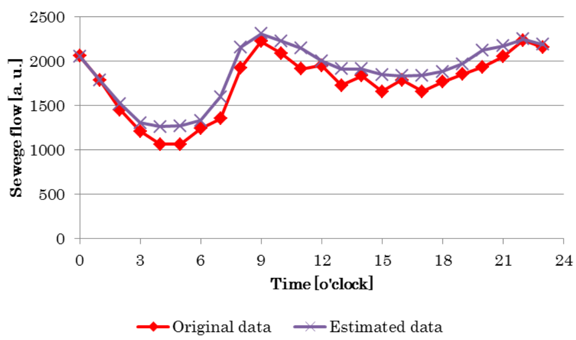


(b) Data with the error around the average one

Fig. 6. Estimation of flow patterns without considering day types.



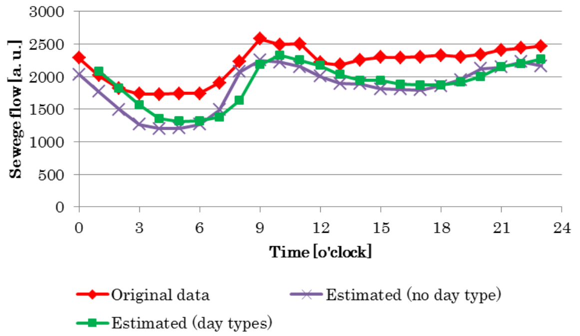
(a) Data with the minimum estimation error



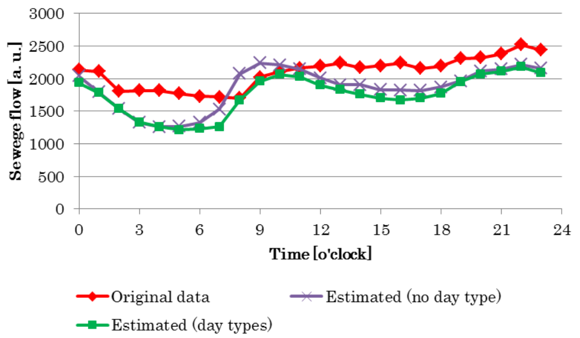
(b) Data with the error around the average one

Fig. 7. Estimation of flow patterns considering day types.

ATLANTIS PRESS International Journal of Networked and Distributed Computing, Vol. 5, No. 4 (October 2017) 233–242



(a) At the largest estimation error in considering no day type



(b) At the largest estimation error in considering day types

Fig. 8. Estimation of flow patterns with large errors.

Results of estimation considering the day type are shown in Fig. 7. Fig. 7 (a) and (b) show the estimation results having the minimum estimation error, and the error around the average one, respectively. The tendency of the estimation considering the day type is quite similar to that without considering it.

Results of estimations with large estimation errors are shown in Fig. 8. The results with the largest estimation error in considering no day type (day types, respectively) are shown in Fig. 8 (a) ((b)). In the figure, both of the results of the estimations with and without considering day types as well as the original values are shown. It can be seen that neither estimation can estimate the original values well.

6. Considerations

The day type is considered to be effective in estimating sewage flow, a kind of daily life data. As daily life data are deeply related to the human activity, it is reasonable that the day type is effective to estimate daily life data.

In estimating sewage flow, the temperature of that day and the amount of the rain fall of the previous day are useful. This means that the temperature directly affects the sewage flow, and the rain fall affects it with time-lag.

Non-regular data are estimated from the typical ones. For this purpose, typical data were firstly tried to be adjusted to non-regular ones. The values used in the adjustment were tried to be estimated by using the multiple regression analysis. Although the regression formulas obtained are for estimating non-regular data, it was experimentally shown that these could be used in estimating regular data. This shows the adaptability and the effectivity of the estimation.

Movements of a body are treated and analyzed.[3-6] The fundamental movements are a kind of time series data. Signal processing methods are used in capturing them. Signal processing methods may be used in representing the sewage flow. The pattern commonly appearing in sewage flows may be clarified.

Representing time series data in the frequency domain may be effective.[7,8] Major information of sewage flow may be obtained by focusing the low frequency components of sewage flow. This kind of representation may derive some important information from the sewage flow.

7. Conclusion

This paper focused on sewage flow as daily life data. We tried to estimate its amounts for their efficient management. Here, we treated non-rainy days. We also treated only the data good enough to process. The day having good data was called a normal day. The normal days were divided into typical days and irregular days. The latter have outliers, and the former do not. The typical days were classified into regular days and non-regular days. The distinction also depends on the existence of outliers. The amounts of flow on regular days are almost those of the typical day.

The amounts of flow of a typical day were tried to be adjusted to those of a non-regular one. The values for the adjustment were tried to be estimated by using the multiple regression analysis. These values could be estimated based on the temperature of that day and the amount of rain fall of the previous day.

This paper also evaluated the degree of the improvement. First, it was shown that the estimation accuracy could be improved, and the estimation errors became smaller than the previous ones. Next, we tried to adapt the estimation of non-regular data from the typical data to regular data. The adaptation was evaluated by using another data set. It was shown that the estimation works well.

There is a room for improving estimation accuracy. More improvement of the estimation is required. This paper treated only the data of non-rainy typical days. Treating irregular days and rainy ones is in future work. The adjustment adopted in this paper does not consider the shift of the peak time of the amounts of flow. Considering the shift of the peak is also included in future work. This paper focused on the amounts of sewage flow. Treating the quality of sewage flow is also in future work.

References

1. Wikipedia, Sewage Treatment, https://en.wikipedia.org/wiki/Sewage_treatment (accessed 2017-03-19)
2. T. Hochin and H. Nomiya, “On the Estimation of Time Series Data of Daily Life,” in *Proc. of 15th IEEE/ACIS International Conference on Software Engineering Research, Management and Applications (SERA 2017)* (2017), pp. 351-356.
3. T. Hochin and H. Nomiya, “Analysis of Pauses Toward Transmitting Traditional Skills,” in *Proc. of 14th IEEE/ACIS International Conferece on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2013)* (2013), pp. 414-419, 2013.
4. T. Hochin and H. Nomiya, “Deriving Fundamental Movements Based on Pauses for Transmitting Traditional Skills,” in *Proc. of 1st ACIS International Symposium on Applied Computers & Information Technology (ACIT 2013)* (2013), pp. 343-348.
5. T. Hochin and H. Nomiya, "Considerations on Pauses and Fundamental Movements for Transmitting Traditional Skills," in *Proc. of 15th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/ Distributed Computing (SNPD 2014)*, (2014), pp.287-292.
6. T. Hochin and H. Nomiya, “Deriving Pauses for Obtaining Fundamental Movements in Traditional Skills,” *Applied Computing and Information Technology, Studies in Computational Intelligence*, 553 (2014) 19-30.
7. T. Hochin, Y. Yamauchi, H. Nakanishi, M. Kojima, H. Nomiya, "Indexing of plasma waveforms for accelerating search and retrieval of their subsequences," *Fusion Engineering and Design*, 85(5) (2010) 649-654.
8. R. Agrawal, C. Faloutsos, and A. N. Swami, "Efficient similarity search in sequence databases," in *Proc. of the 4th International Conference on Foundations of Data Organization and Algorithms (FODO'93)* (1993), pp. 69-84.