

Project Maintenance Effort Optimization Based on Flexible JDP Model for OSS Fault Big Data

Yoshinobu Tamura

Department of Intelligent Systems,
Tokyo City University,
Tamazutsumi 1-28-1, Setagaya-ku, Tokyo 158-8557, Japan.
Corresponding author: tamuray@tcu.ac.jp

Shigeru Yamada

Graduate School of Engineering,
Tottori University,
Minami 4-101, Koyama, Tottori-Shi, 680-8552, Japan.
E-mail: yamada@tottori-u.ac.jp

(Received March 22, 2019; Accepted September 9, 2019)

Abstract

We focus on double irregular fluctuations under jump in the operation performance of open source software (OSS). Then, this paper proposes the method of cost optimization based on flexible jump diffusion process (JDP) model in order to consider several noisy cases for maintenance effort in the OSS operation with version upgrade. In particular, we discuss a method of effort optimization based on the flexible JDP model with the unexpected irregular continuous fluctuation in version upgrade for OSS projects. The proposed method will be useful for the OSS project managers to decide the optimal version upgrade and maintenance time of OSS under the OSS project management. Furthermore, we show several analysis examples of the optimization method considering the properties of version upgrade under OSS projects.

Keywords-Fault big data, Jump diffusion process, Software effort, Effort optimization.

1. Introduction

A large number of open source software (OSS) have been developed around the world such as OpenStack, Firefox, Apache HTTP server, OpenOffice, and PostgreSQL, etc. Many OSS include various versions known as bug-fix version, minor version, and major version. Considering the OSS management projects, the system performance of OSS management will take an unexpected irregular continuous fluctuation in the specified terms of operation under specific version, because many users become committed to the OSS maintenance in specific version. In particular, the styles of management and usage take on different forms depending on each version of OSS. Several research papers for OSS reliability assessment have been published by specified researchers (Norris, 2004; Yamada and Tamura, 2016). However, several research papers discussed the effort expenditures of OSS project have not been proposed. At that time, many research papers have been focused on the software reliability by analyzing the fault data. For example, the analysis methods of software reliability by using the classical software reliability growth models have been discussed by several researchers (Kapur et al., 2011; Yamada, 2014). Moreover, our OSS research team has been proposed the analysis method of OSS reliability (Yamada and Tamura, 2016). In particular, it is very important for the project managers to supervise the OSS maintenance effort for quality management. The proper control of OSS maintenance effort has an indirect influence on the reliability, cost, and quality. Furthermore, the proper control of OSS effort indirectly decrease

the software faults. Therefore, it will make an improvement the OSS quality if the control effort of OSS management is appropriately controlled under the operation.

This paper discusses a stochastic model for OSS effort management with unexpected irregular continuous fluctuation in performance from the characteristics with version upgrade of OSS development. This paper proposes the method of effort optimization based on the flexible jump diffusion process (JDP) models in terms of effort in order to consider the properties with version upgrade of OSS. Then, we formulate the optimal OSS maintenance effort expense problem based on the discussed flexible JDP models. Then, we determine the optimum OSS maintenance time minimized the expected total OSS maintenance effort. In particular, we discuss the progress stability of OSS project under the specified version by using the estimated sample path of expected total OSS effort based on the flexible JDP model. Furthermore, this paper shows the analysis examples for the proposed method of OSS maintenance effort optimization using flexible JDP models.

2. Flexible JDP Model

We discuss a JDP model to control the OSS maintenance effort in the operation. Let $\Psi(t)$ be the cumulative OSS maintenance effort expenditures up to time t ($t \geq 0$) in the operation. $\Psi(t)$ is the real values continuously. Then, $\Psi(t)$ gradually increases as the progress of OSS operation, because the estimated maintenance effort are recorded during the OSS operation. By using the modeling technique of classical software reliability growth models, the following typical linear differential equation considering the OSS maintenance effort is given as follows:

$$\frac{d\Psi(t)}{dt} = \beta(t)\{\alpha - \Psi(t)\}, \quad (1)$$

where $\beta(t)$ is the effort expenditure rate of OSS at time t and a non-negative function, α , means the estimated OSS maintenance effort expended until the end of specified version.

Considering Brownian motion, Eq. (1) is extended to the following stochastic differential equation (SDE) (Arnold, 1974; Yamada et al., 1994):

$$\frac{d\Psi(t)}{dt} = \{\beta(t) + \sigma v(t)\}\{\alpha - \Psi(t)\}, \quad (2)$$

where σ is added as a positive constant value representing a magnitude of the irregular continuous fluctuation, and $v(t)$ a standardized Gaussian white noise due to development environment. Then, Eq. (2) is extended to the following SDE based on an Itô type:

$$d\Psi(t) = \left\{\beta(t) - \frac{1}{2}\sigma^2\right\}\{\alpha - \Psi(t)\}dt + \sigma\{\alpha - \Psi(t)\}dw(t), \quad (3)$$

where $w(t)$ means one-dimensional Wiener process. $w(t)$ can be formally defined as an integration of the white noise $v(t)$.

The jump term is embedded to the SDE models of Eq. (3) in order to consider the unexpected irregular state around time t by many external complicated factors in the OSS project. The JDP (Merton, 1976) is obtained as follows:

$$d\Psi_j(t) = \left\{ \beta(t) - \frac{1}{2}\sigma^2 \right\} \{ \alpha - \Psi_j(t) \} dt + \sigma \{ \alpha - \Psi_j(t) \} dw(t) + d \left\{ \sum_{i=1}^{Y_t(\lambda)} (V_i - 1) \right\}, \quad (4)$$

where $Y_t(\lambda)$ means a Poisson point process with frequency λ at time t . Also, $Y_t(\lambda)$ is the number of jumps, and λ the rate of jump. V_i means the range of i -th jump. We assume that $w(t)$, $Y_t(\lambda)$, and V_i are independent mutually. Moreover, we define the increase rate of OSS maintenance effort for $\beta(t)$ as:

$$\int_0^t \beta(s) ds \doteq \frac{dR_s(t)}{\alpha - R_s(t)}, \quad (5)$$

$$R_e(t) = a(1 - e^{-bt}), \quad (6)$$

$$R_s(t) = a\{1 - (1 + bt)e^{-bt}\}. \quad (7)$$

In this paper, $\beta(t)$ is assumed the mean value functions in Eqs. (6) and (7) based on software reliability models as the OSS effort expenditure function of the proposed model, where $a \doteq \alpha$ is the expected number of latent faults, and $b \doteq \beta$ the fault detection rate per fault in terms of software reliability growth models.

Based on Itô's formula (Arnold, 1974; Yamada et al., 1994), the solution of Eq. (4) can be derived as follows:

$$\Psi_{je}(t) = \alpha \left[1 - \exp \left\{ -\beta t - \sigma w(t) - \sum_{i=1}^{Y_t(\lambda)} \log V_i \right\} \right], \quad (8)$$

$$\Psi_{js}(t) = \alpha \left[1 - (1 + \beta t) \exp \left\{ -\beta t - \sigma v(t) - \sum_{i=1}^{Y_t(\lambda)} \log V_i \right\} \right]. \quad (9)$$

Moreover, we extend the existing JDP model obtained from Eq. (4) to the following time-delay JDPes:

In case of $(t \geq 0)$:

$$d\Psi_{fj}(t) = \left\{ \beta(t) - \frac{1}{2}\sigma^2 \right\} \{ \alpha - \Psi_{fj}(t) \} dt + \sigma \{ \alpha - \Psi_{fj}(t) \} dw(t) + d \left\{ \sum_{i=0}^{Y_t(\lambda_1)} (V_i^1 - 1) \right\} \quad (10)$$

In case of $(t \geq 0, t' \geq t_1)$:

$$d\Psi_{fj}(t) = \left\{ \beta(t) - \frac{1}{2}\sigma^2 \right\} \{ \alpha - \Psi_{fj}(t) \} dt + \sigma \{ \alpha - \Psi_{fj}(t) \} dw(t) \\ + d \left\{ \sum_{i=0}^{Y_t(\lambda_1)} (V_i^1 - 1) \right\} + d \left\{ \sum_{i=0}^{Y_{t'}(\lambda_2)} (V_i^2 - 1) \right\} \quad (11)$$

where $Y_t(\lambda_1)$ and $Y_{t'}(\lambda_2)$ are Poisson point processes with parameter λ_1 and λ_2 at each operation time $(t \geq 0)$ and $(t' \geq t_1)$, respectively. Moreover, V_i^1 and V_i^2 are i -th jump ranges in each operation time $(t \geq 0)$ and $(t' \geq t_1)$, respectively. We assume that $Y_t(\lambda_1)$, $Y_{t'}(\lambda_2)$, V_i^1 , and V_i^2 are mutually independent in this paper.

From Itô's formula (Arnold, 1974; Yamada et al., 1994), the solution of Eqs. (10) and (11) can be obtained as follows:

In case of ($t \geq 0$):

$$\Psi_{fje}(t) = \alpha \left[1 - \exp \left\{ -\beta t - \sigma w(t) - \sum_{i=1}^{Y_t(\lambda_1)} \log V_i^1 \right\} \right] \quad (12)$$

$$\Psi_{fjs}(t) = \alpha \left[1 - (1 + \beta t) \cdot \exp \left\{ -\beta t - \sigma w(t) - \sum_{i=1}^{Y_t(\lambda_1)} \log V_i^1 \right\} \right] \quad (13)$$

In case of ($t \geq 0, t' \geq t_1$):

$$\Psi_{fje}(t) = \alpha \left[1 - \exp \left\{ -\beta t - \sigma w(t) - \sum_{i=1}^{Y_t(\lambda_1)} \log V_i^1 - \sum_{i=1}^{Y_{t'}(\lambda_2)} \log V_i^2 \right\} \right] \quad (14)$$

$$\Psi_{fjs}(t) = \alpha \left[1 - (1 + \beta t) \cdot \exp \left\{ -\beta t - \sigma w(t) - \sum_{i=1}^{Y_t(\lambda_1)} \log V_i^1 - \sum_{i=1}^{Y_{t'}(\lambda_2)} \log V_i^2 \right\} \right] \quad (15)$$

Considering the time-delay over t_2 ($t_2 \geq t_1$), we can formulate the flexible JDP models as follows:

$$\Psi_{fje}(t) = \alpha \left[1 - \exp \left\{ -\beta t - \sigma w(t) - \sum_{k=1}^K \sum_{i=1}^{Y_{t^k}(\lambda_k)} \log V_i^k \right\} \right], \quad (16)$$

$$\Psi_{fjs}(t) = \alpha \left[1 - (1 + \beta t) \exp \left\{ -\beta t - \sigma w(t) - \sum_{k=1}^K \sum_{i=1}^{Y_{t^k}(\lambda_k)} \log V_i^k \right\} \right], \quad (17)$$

where t^k ($k = 1, 2, \dots, K$) means k -th specific time for major version upgrade, and K is the number of major version upgrade.

3. Parameter Estimation for Flexible JDP Models

We use the method of maximum-likelihood for the unknown parameters α , β , b , and σ . In particular, it is difficult to make a decision the jump parameters of JDP models because of the complexity in likelihood function included multiple distributions based on the Wiener process and jump diffusion one. Several estimation methods for jump parameters of JDP model have been proposed by the specified researchers. However, there are no effective methods of such an estimation. This paper discusses the estimation method of parameters in terms of jump terms. Then, a genetic algorithm (GA) is used in this paper (Holland, 1975; Honoré, 1998; Tamura and Yamada, 2015) in order to make a decision the jump parameters of the discussed model.

For example, we assume that our JDP models includes the parameters λ_1 and λ_2 for Y_t and $Y_{t'}$, similarly, μ_1 , μ_2 , τ_1 , and τ_2 for V_i^1 and V_i^2 in Eqs.(11) and (12). Then, the set parameters J in terms of λ_1 , μ_1 , and τ_1 are estimated by GA algorithm in case of ($t \geq 0$). Similarly, the set parameters J' in terms of λ_2 , μ_2 , and τ_2 make a decision by using the GA algorithm in case of ($t' \geq t_1$).

Exponential type model in case of ($t \geq 0$):

$$\begin{cases} \min_{J_e} \Psi_i^e(J_e) \\ \Psi_i^e = \sum_{i=0}^K \{\Psi_{fje}(i) - x_i\}^2, \end{cases} \quad (18)$$

S-shaped type model in case of ($t \geq 0$):

$$\begin{cases} \min_{J_s} \Psi_i^s(J_s) \\ \Psi_i^s = \sum_{i=0}^K \{\Psi_{fjs}(i) - x_i\}^2, \end{cases} \quad (19)$$

Exponential type model in case of ($t' \geq t_1$):

$$\begin{cases} \min_{J_{e'}} \Psi_i^e(J_{e'}) \\ \Psi_i^e = \sum_{i=0}^K \{\Psi_{fje}(i) - x_i\}^2, \end{cases} \quad (20)$$

S-shaped type model in case of ($t' \geq t_1$):

$$\begin{cases} \min_{J_{s'}} \Psi_i^s(J_{s'}) \\ \Psi_i^s = \sum_{i=0}^K \{\Psi_{fjs}(i) - x_i\}^2, \end{cases} \quad (21)$$

where $\Psi_{fje}(i)$ and $\Psi_{fjs}(i)$ are the cumulative amount of OSS maintenance effort at time i in the JDP model, x_i the actual project data in terms of the cumulative amount of OSS maintenance effort.

4. Maintenance Effort Optimization

As with the classical optimal software release problems (Yamada and Osaki, 1985; Yamada and Osaki, 1987), the effort parameters are defined as follows:

- ρ_1 : the importance rate of OSS maintenance effort expended during the operation,
- ρ_2 : the rate of OSS operation effort expenditures per unit time,
- ρ_3 : the importance rate of OSS operation effort after the upgrade.

The expected OSS effort expenditures in the software operation phase can be formulated by:

$$M_1(t) = \rho_1 E[\Psi(t)] + \rho_2 t. \quad (22)$$

The expected OSS maintenance effort expenditures after the version upgrade of OSS is formulated as follows:

$$M_2(t) = \rho_3\{\alpha - E[\Psi(t)]\}. \quad (23)$$

Finally, using Equations (22) and (23), the total OSS effort is represented by

$$M(t) = M_1(t) + M_2(t). \quad (24)$$

As for the sample path for several jump terms, we can consider as the following equations. The sample path of exponential type model is given as

$$M_{fje}(t) = \rho_1\Psi_{fje}(t) + \rho_2t + \rho_3\{\alpha - \Psi_{fje}(t)\}. \quad (25)$$

Similarly, the sample path of S-shaped type model is given as

$$M_{fjs}(t) = \rho_1\Psi_{fjs}(t) + \rho_2t + \rho_3\{\alpha - \Psi_{fjs}(t)\}. \quad (26)$$

5. Numerical Examples

We discuss the actual data in Apache HTTP Server (The Apache Software Foundation). Apache HTTP Server is developed under Apache Software Foundation as OSS. The actual data set is shown in Figure 1. Figure 2 is the estimated cumulative effort for OSS operation expenditures in case of $E_{fje}(t)$. In Figure 2, the long-dash line shows the start line of version 4.1.31 major-version-upgraded from version 3.x line. Also, the dot-dash line shows the start line of beta version 7.0.0 major-version-upgraded from version 6.x line. In Figure 2, we find that the noises based on unexpected jumps become large after the operation time 1,826 days. Moreover, the estimated OSS maintenance effort for $E_{fjs}(t)$ is shown in Figure 3. From Figure 3, we find that the $E_{fjs}(t)$ fits better than $E_{fje}(t)$ for the practical effort data.

Moreover, Figure 4 shows the estimated sample path of total OSS effort expenditures for exponential type model. Similarly, the estimated sample path of total OSS effort in case of S-shaped type model is shown in Figure 5. From Figure 4, we found that the optimum OSS maintenance time is derived as $t^* = 13.635$ years. This result shows that there is need to make a transition to major upgrade after 13.6 years from the release of specified version. Similarly, Figure 5 means that the optimum OSS maintenance time is estimated as $t^* = 17.314$ years. In particular, the noise becomes large after 4,000 days. Therefore, the software manager has to move up the optimum maintenance time from t^* to 4,000 days (11 years) in terms of the noise. Thus, the proposed method can support the stability by using several noises. Moreover, it will be useful to control the maintenance effort considering the OSS operational environment.

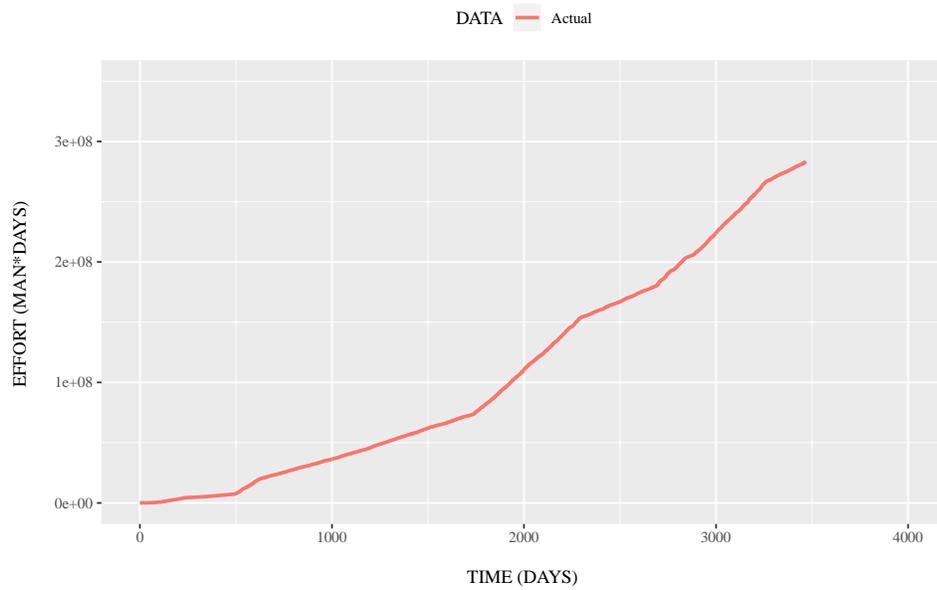


Figure 1. The actual effort data set

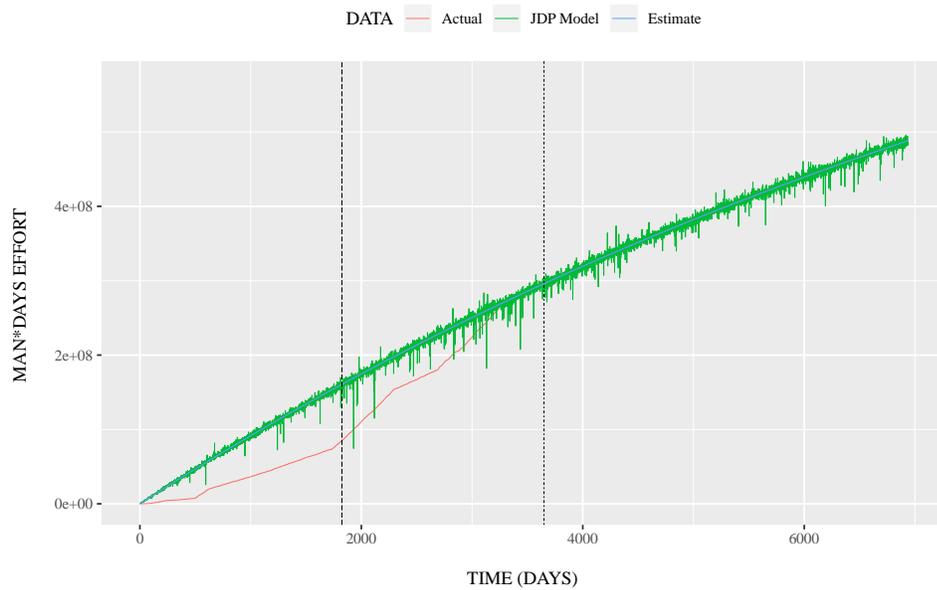


Figure 2. The estimated cumulative OSS maintenance effort expenditures for $E_{fje}(t)$

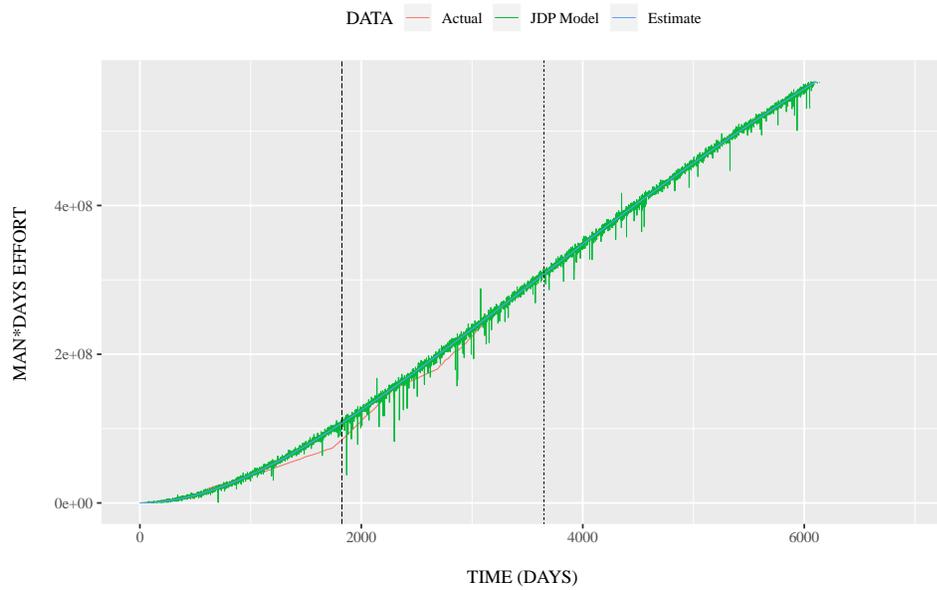


Figure 3. The estimated cumulative OSS maintenance effort expenditures for $E_{fjs}(t)$

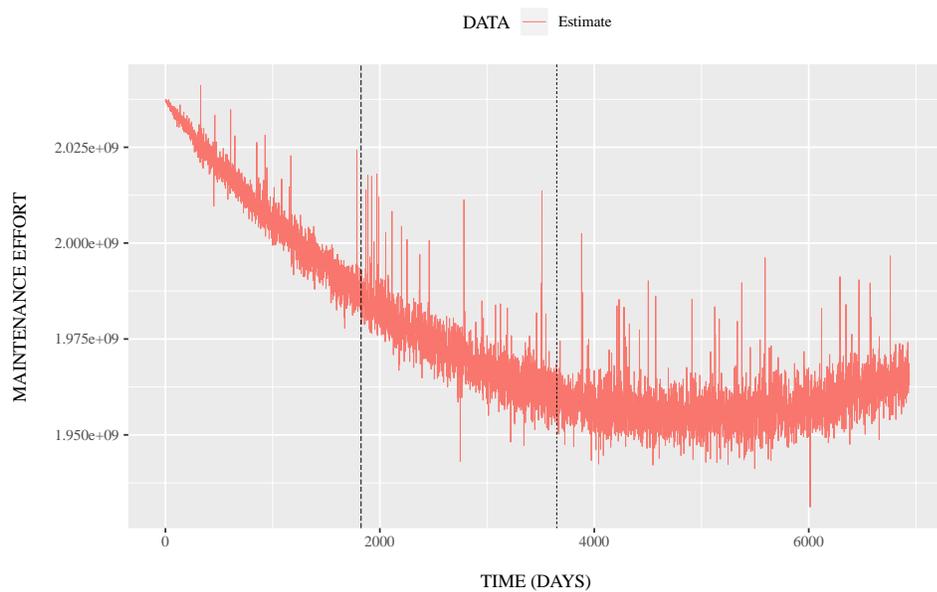


Figure 4. The sample path of estimated total OSS effort for exponential type model

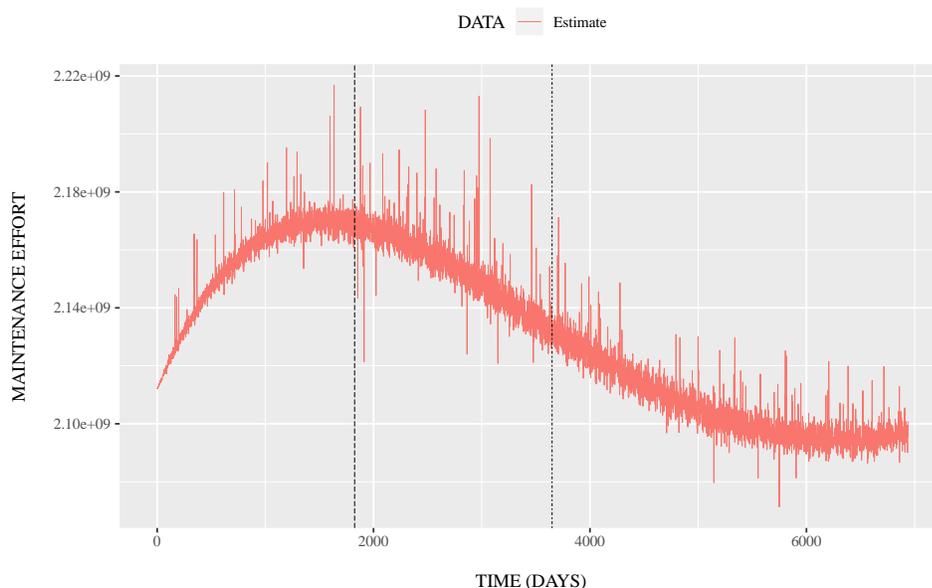


Figure 5. The sample path of estimated total OSS effort for S-shaped type model

6. Concluding Remarks

Various versions are included in many OSS. There are many types of version such as bug-fix version, minor version, and major version. The proper control of OSS maintenance effort will indirectly relate to the cost and reliability of OSS. However, it will be difficult for the OSS managers and users to make a progress management of OSS project considering the periods of specified version according to several version upgrade. Considering the characteristics included several version upgrade of OSS development and management, this paper has discussed the method of OSS maintenance effort optimization by using the flexible JDP model with the unexpected irregular continuous fluctuation with jump. This method will be useful as the determination method of the proper progress for version upgrade in OSS operation. Also, the optimal OSS maintenance effort expense problem with OSS project stability in order to decide the optimal time has been discussed in Section 4. In particular, the amount of noise for the total maintenance effort will be helpful to make a decision the optimal time with progress stability of OSS project.

Conflict of Interest

The authors confirm that there is no conflict of interest to declare for this publication.

Acknowledgements

The authors would like to express their sincere thanks to the editor and anonymous reviews for their time and valuable suggestions.

References

- Arnold, L. (1974). *Stochastic differential equations—theory and applications*. John Wiley & Sons, New York.
- Holland, J.H. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press. USA.
- Honoré, P. (1998). *Pitfalls in estimating jump-diffusion models*. Working Paper Series 18, University of Aarhus, School of Business.
- Kapur, P.K., Pham, H., Gupta, A., & Jha, P.C. (2011). *Software reliability assessment with OR applications*. (p. 364). London: Springer.
- Merton, R.C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2), 125-144.
- Norris, J.S. (2004). Mission-critical development with open source software: lessons learned. *IEEE Software*, 21(1), 42-49.
- Tamura, Y., & Yamada, S. (2015). Reliability analysis based on a jump diffusion model with two Wiener processes for cloud computing with big data. *Entropy*, 17(7), 4533-4546.
- Yamada, S. (2014). *Software reliability modeling: fundamentals and applications*, (Vol. 5) Springer-Verlag, Tokyo/Heidelberg.
- Yamada, S. and Tamura, Y. (2016). *OSS reliability measurement and assessment*. Springer International Publishing, Switzerland.
- Yamada, S., & Osaki, S. (1985). Cost-reliability optimal software release policies for software systems. *IEEE Transactions on Reliability*, R-34(5), 422-424.
- Yamada, S., & Osaki, S. (1987). Optimal software release policies with simultaneous cost and reliability requirements. *European Journal of Operational Research*, 31(1), 46-51.
- Yamada, S., Kimura, M., Tanaka, H., & Osaki, S. (1994). Software reliability measurement and assessment with stochastic differential equations. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 77(1), 109-116.

Link and Websites

The Apache Software Foundation. The Apache HTTP Server Project, <http://httpd.apache.org/>



Original content of this work is copyright © International Journal of Mathematical, Engineering and Management Sciences. Uses under the Creative Commons Attribution 4.0 International (CC BY 4.0) license at <https://creativecommons.org/licenses/by/4.0/>