

## 2変量ワイブル型ソフトウェア信頼度成長モデルとその適合性評価

井上 真二<sup>†1</sup> 山田 茂<sup>†1</sup>

本論文では、ワイブル型ソフトウェア信頼度成長モデル (SRGM) の信頼度成長影響要因としてのテスト時間をコブ・ダグラス型効用関数の考え方にに基づきながら2次元時間空間へと拡張することによって、ソフトウェア信頼性評価のための2変量ワイブル型SRGMを構築する。また、提案モデルのパラメータ推定手法やソフトウェア信頼性評価尺度について議論するとともに、組み込みソフトウェアのテスト工程で得られた実測データを用いながら提案モデルの適用例も示す。最後に、代表的な1変量SRGMを取り上げ、実測データを用いた適合性比較を行いながら、提案モデルの信頼性評価精度に関する性能を検証して、その理論的および実験的考察を与える。

### A Bivariate Weibull-type Software Reliability Growth Model and Its Goodness-of-fit Evaluation

SHINJI INOUE<sup>†1</sup> and SHIGERU YAMADA<sup>†1</sup>

We propose a bivariate software reliability growth model (SRGM) describing a software reliability growth process on the two-dimensional time space, which consists of the testing-time and the other reliability growth factors. Especially, this paper extend the testing-time, which is regarded as a reliability growth factor in a conventional (one-dimensional) SRGM, to the two-dimensional time space by applying the Cobb-Douglas utility function. And we also discuss a parameter estimation method and a software reliability assessment measure of our bivariate model. Further, we show numerical examples on software reliability analysis based on our model by using fault count data collected from actual testing phases of embedded software systems. Finally, comparing with representative one-dimensional SRGMs, we investigate the performance on software reliability assessment accuracy of our model by using actual data sets.

<sup>†1</sup> 鳥取大学大学院工学研究科社会基盤工学専攻

Department of Management of Social Systems and Civil Engineering, Graduate School of Engineering, Tottori University

### 1. はじめに

高信頼性を有したソフトウェアをユーザに提供するためには、ソフトウェア開発工程の上流工程におけるレビュー活動などの品質管理活動とともに、その最終工程であるテスト工程におけるソフトウェア品質/信頼性の十分な計測・評価が重要である。テスト工程におけるソフトウェア信頼性評価技術の中で、ソフトウェア信頼度成長モデル (software reliability growth model, 以下SRGMと略す)<sup>1)-6)</sup>は、現在、定量的なソフトウェア信頼性評価のための基盤技術の1つとして浸透している。具体的に、SRGMとは、テスト期間中に観測される発見フォールト数やソフトウェア故障発生時間間隔に関するデータに基づいて表現されるソフトウェア信頼度成長過程を確率・統計則に基づいて記述する数理モデルであり、フォールト発見数モデルとソフトウェア故障発生時間モデルに大別される<sup>3)</sup>。特に、フォールト発見数モデルは、評価に必要なデータ(フォールト発見数データ)がソフトウェア故障発生時間モデルに必要なデータ(ソフトウェア故障発生(間隔)時間データ)に比べて低労力で得られるため、多く実用に供されているモデルである。これまでに提案されているフォールト発見数モデルの多くは、基本的に、ソフトウェア信頼度成長過程をテスト時間のみに依存した確率過程として記述している。しかしながら、ソフトウェア信頼度成長過程がテスト時間のみに依存するという基本的仮定に基づいて構築されたSRGMは、実質的なテスト作業時間やテスト十分性などソフトウェア故障発生現象に大きく影響を与える要因は加味せず、テストをできるだけ長く行えば、修正・除去されるフォールト数は増加し、それに応じてソフトウェアの信頼性は向上することを暗に前提としている。

一般的に、ソフトウェア故障は、ソフトウェアフォールトが内在するプログラムパスが、入力データ(テストケース)によって実行された際に発生する。したがって、このようなソフトウェア故障発生メカニズムや実際のテスト工程におけるデバッグ作業を考えた場合、たとえテストを長時間実施したとしても、それに応じたCPU時間などの実質的なテスト作業時間やテスト十分性がともなっていないければ、作り込まれたフォールトの存在を顕在化することはできず、ソフトウェアの信頼性を向上させることもできない。また、現実の問題として、CPU時間やテスト十分性などの様々な要因が、テスト工程におけるソフトウェア信頼度成長過程に大きく影響を与えるということが数多く報告されており、それらの状況を考慮したSRGMも数多く提案されている<sup>7)-9)</sup>。ただし、ソフトウェア信頼度成長影響要因を考慮したこれらのSRGM<sup>7)-9)</sup>に従って記述されるソフトウェア信頼度成長過程は、本質的にはテスト時間のみに依存したものであり、信頼度成長過程のとらえ方という観点から

は、実用に多く供されている指数形 SRGM<sup>10)</sup> や遅延 S 字形 SRGM<sup>11)</sup> などのモデルと比較しても、その本質的な違いが見受けられない。

これらの背景に鑑みると、実際に観測されるソフトウェア信頼度成長現象をより現実的にモデルへと反映させるためには、テスト時間とともに、上述した CPU 時間やテスト十分ななどの信頼度成長要因がソフトウェア信頼度成長過程に与える影響も同時に考慮する必要がある。近年、その解決策の 1 つとして、多次元時間空間上においてソフトウェア信頼度成長過程を記述するモデルの開発が行われている。文献 12) では、有限フォールト数を仮定した 1 次元非同次ポアソン過程 (one-dimensional nonhomogeneous Poisson process, 以下 1 次元 NHPP と略す) モデルに対する既存の一般化枠組みに基づき、そのソフトウェア故障発生時間分布が 2 次元時間空間上で表現される場合を考えることによって、2 変量ソフトウェア信頼性モデルの構築枠組みを提案している。この枠組みの特徴は、テスト時間 (カレンダー時間) とテスト労力量 (テスト実行時間) から構成される 2 変量ソフトウェア故障発生時間分布を当該枠組みに対して仮定することで、検出可能フォールト数が有限個であると仮定された様々なタイプの 2 変量 SRGM を構築することができる点にある。ただし、当該枠組みにおいてテスト実行時間以外のものをテスト労力要因を表す指標として利用する場合、若干の注意が必要である。すなわち、当該枠組みでは、テスト時間とテスト実行時間がそれぞれ非負の確率変数として表現されており、テスト網羅度 (testing-coverage) など、とりうる値の範囲がそもそも決まっている指標をテスト労力量として利用した場合、当該枠組みにおける 2 変量ソフトウェア故障発生時間分布が通常の 2 変量確率分布として適切に機能しないことが考えられる。テスト網羅度は、テストケースがプログラムパスに対して網羅的かつ効率的に設計されているかどうかを表す指標であり、ソフトウェア信頼度成長過程に影響を与える重要な要因の 1 つである。特に、大規模化・複雑化が進む組み込みソフトウェアのテストでは、テスト網羅度は重要視されており、網羅度達成状況とともにそれとともない観測される信頼度成長現象を同時に把握できるモデルを開発することは、昨今の重要な課題であるといっても過言ではない。

本論文では、上述した背景をふまえ、文献 12) とは異なったアプローチによって、既存の 1 変量 SRGM の時間空間を 2 次元時間空間へと拡張し、2 変量 SRGM を構築する。具体的には、既存の 1 変量 SRGM におけるソフトウェア信頼度成長要因としてのテスト時間を、テスト時間要因とテスト労力要因からなる一種の抽象的要因であると考え、それらをコブ・ダグラス効用関数の考え方をを用いて表現する。特に、本論文では、検出可能フォールト数が無限の場合を仮定して構築される既存のワイブル型ソフトウェア信頼度成長モデル

(Weibull-type SRGM, 以下ワイブル型モデルと略記する) を取り上げ、上述した考え方に基づきながら、ソフトウェア信頼性評価のための 2 変量ワイブル型モデル (bivariate Weibull-type SRGM) を構築する。その後、提案モデルのパラメータ推定手法および提案モデルに基づくソフトウェア信頼性評価尺度について議論する。また、実測データを用いながら提案モデルに基づいたソフトウェア信頼性解析例を示すと同時に、提案モデルと既存の代表的な 1 変量 SRGM との実測データに対する適合性比較を行い、提案モデルの信頼性評価精度に関する性能を検証する。今回議論するモデリングアプローチの本質的内容に関しては、著者が以前、研究報告<sup>13)</sup> として公表しているが、本論文では、実際の組み込みソフトウェアシステムのテストにおいて得られた 2 組の実測データを用いながら、様々な評価基準に基づいた提案モデルの適合性評価および既存モデルとの性能比較を行い、さらに、これらの適合性評価結果に関するいくつかの考察を与えることで、上記研究報告において公表した内容も含めながら提案モデルの実用性および有効性について包括的に議論する。

## 2. ワイブル型モデル

本論文において提案する 2 変量ワイブル型モデルは、既存の 1 変量ワイブル型モデル<sup>2),3),14)</sup> の時間空間を 2 次元へ拡張することによって導出される。いま、 $\{N(t), t \geq 0\}$  を、任意のテスト時刻  $t$  までに発見された総フォールト数を表す確率過程とする。 $N(t)$  が平均値関数 (mean value function)  $H(t)$  の NHPP に従うものと仮定した場合、 $\{N(t), t \geq 0\}$  の確率法則は、

$$\left. \begin{aligned} \Pr\{N(t) = n\} &= \frac{\{H(t)\}^n}{n!} \exp[-H(t)] \\ &\quad (n = 0, 1, 2, \dots) \\ H(t) &= \int_0^t h(\tau) d\tau \end{aligned} \right\} \quad (1)$$

のように与えられる。ここで、 $\Pr\{A\}$  は事象  $A$  の確率、 $h(t)$  は NHPP の強度関数 (intensity function) をそれぞれ表す。式 (1) の NHPP は、その適用性および実用性の観点から、フォールト発見事象もしくはソフトウェア故障発生現象を記述する際に多く用いられている確率過程の 1 つである。フォールト発見事象の確率的挙動に関して、式 (1) の NHPP を仮定した場合、実際にソフトウェア信頼性評価を行うためには、平均値関数  $H(t)$  を具体的に与える必要がある。これまでに、実際のテスト環境やソフトウェア信頼度成長過程に影響を与える要因などを考慮しながら、数多くの平均値関数が構築されている。ここで、平均値関

数  $H(t)$  が,

$$H(t) \equiv \gamma(t) = \left(\frac{t}{\rho}\right)^\beta \quad (0 < \beta < 1; \rho > 0) \quad (2)$$

のように与えられた場合, 式 (1) および式 (2) からなる NHPP モデルは, ワイブル型モデル<sup>2),3)</sup> と呼ばれる. ここで,  $\beta$  は信頼度成長パラメータ,  $\rho$  は尺度パラメータである. 式 (2) は, その関数の構造上,  $\gamma(0) = 0$  および  $\gamma(\infty) = \infty$  となる. これは, 無限に長くテスト作業を実施すれば, そのテスト作業によって無限個のフォールトが検出されることを意味する. このような現象は, テスト工程におけるデバッグ作業が必ずしも完全ではなく, それによって新たに作り込まれるフォールトの存在を現実的な仮定として認めた場合に考えられる現象であり, 特に, そのような性質を持つ SRGM は, 一般的に, 不完全デバッグモデル (imperfect debugging model)<sup>3),4)</sup> と呼ばれている.

### 3. 2変量ワイブル型モデルの構築

本論文では, 従来のワイブル型モデルにおけるソフトウェア信頼度成長要因としてのテスト時間を以下の2つの要因に大別する:

- テスト時間要因 ( $s$ ),
- テスト労力要因 ( $u$ ).

本論文において, 上述の「テスト時間要因」とはテスト開始時刻からのカレンダー時間を意味する. また, 「テスト労力要因」とは実質的なテスト実行時間 (CPU 時間), 実行されたテストケース数, 達成されたテスト網羅などを指しており, テスト時間要因以外の信頼度成長要因の中でも実際のテスト工程において計測可能であり, テストケース設計者などのテスト作業者がフォールトを検出するために費やした実質的なテスト労力を意味する. この分類は, フォールト発見数データが一般的にカレンダー時間に沿って収集されることと, カレンダー時間に沿って長期間テストを実施したとしても実質的に信頼度成長過程に影響を与える要因がともなっていないと信頼性は向上しないという現実的なソフトウェア信頼度成長現象をモデルにおいて表現するためである. なお, 従来の1変量 SRGM では, カレンダー時間のみが1次元時間空間の要素として主に取り扱われていることに注意されたい.

本論文では, 上述した2つのソフトウェア信頼度成長要因に基づいて, コブ・ダグラス効用関数 (Cobb-Douglas utility function)<sup>15),16)</sup> の考え方を導入する. すなわち, 従来のワイブル型モデルにおけるソフトウェア信頼度成長要因としてのテスト時間  $t$  を,

$$t \equiv s^\alpha u^{1-\alpha} \quad (0 \leq \alpha \leq 1) \quad (3)$$

として表現する. ここで,  $\alpha$  はソフトウェア信頼度成長要因としての影響度合いを表すパラメータである. 続いて, 上述した概念に基づいた SRGM を構築するために, 任意のテスト時刻 ( $s, u$ ) までに検出された総フォールト数を表す2次元確率過程  $\{N(s, u), s \geq 0, u \geq 0\}$  (two-dimensional stochastic process)<sup>17),18)</sup> を導入する. 本論文では, 2次元確率過程  $N(s, u)$  が以下に示す2次元 NHPP (two-dimensional NHPP) に従うものと仮定する.

$$\left. \begin{aligned} \Pr\{N(s, u) = n\} &= \frac{\{H(s, u)\}^n}{n!} \exp[-H(s, u)] \quad (n = 0, 1, 2, \dots) \\ H(s, u) &= \int_0^s \int_0^u h(\tau, v) d\tau dv \end{aligned} \right\} \quad (4)$$

ここで,  $H(s, u)$  および  $h(s, u)$  は, それぞれ, 2次元 NHPP の平均値関数および強度関数を表す. 本論文では, 式 (4) の平均値関数  $H(s, u)$  に関して, 式 (2) の1次元テスト時間空間を式 (3) の概念に基づいて2次元時間空間へ拡張した関数:

$$H(s, u) \equiv \gamma(s, u) = \left(\frac{s^\alpha u^{1-\alpha}}{\rho}\right)^\beta \quad (5)$$

を用いる. 式 (5) において,  $\alpha = 0$  のとき,  $t = u$  となり, テスト労力のみ依存するテスト労力依存型 SRGM (Testing-effort dependent SRGM)<sup>19)</sup> となる. また,  $\alpha = 1$  のとき  $t = s$  となり, 従来の1変量ワイブル型モデルと本質的に等価となる.

### 4. ソフトウェア信頼性評価尺度

3章において議論した2変量ワイブル型モデルに基づき, 定量的なソフトウェア信頼性評価に有用な信頼性評価尺度について議論する. 本論文では, ソフトウェア運用信頼度 (operational software reliability)<sup>12)</sup> について議論する.

ソフトウェア運用信頼度とは, 評価開始時点がリリース (出荷) 時期と同じであり, 以後, テスト労力がまったく投入されない場合を仮定したときのソフトウェア信頼度を意味する. すなわち, テストが終了時刻  $s_e$  まで進行しており, テスト開始時刻からその時点までの総テスト労力投入量が  $u_e$  であるとき, 以後の (運用) 時間区間  $(s_e, s_e + \eta]$  ( $s_e \geq 0, \eta \geq 0$ ) においてソフトウェア故障が発生しない確率を表す. 図1に, ソフトウェア運用信頼度の概念図を示す. これより, ソフトウェア運用信頼度は, 式 (4) および式 (5) より,

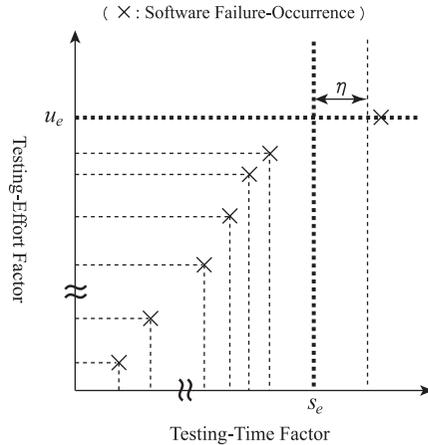


図1 ソフトウェア運用信頼度の概念図  
Fig.1 Basic concept of operational software reliability.

$$R(\eta | s_e, u_e) = \exp \left\{ - \left[ \gamma(s_e + \eta, u_e | \hat{\Theta}) - \gamma(s_e, u_e | \hat{\Theta}) \right] \right\} \quad (6)$$

と導出される．ここで， $\hat{\Theta}$  は，式 (5) に含まれる推定されたパラメータ全体を表す．

### 5. パラメータ推定

提案したモデルに含まれるパラメータ  $\alpha, \beta$ ，および  $\rho$  を重回帰分析を用いて推定する．ここで，一定のテスト時間区間  $(0, s_k]$  において，投入された総テスト労力量  $u_k$  および発見された総フォールト数  $y_k$  に関する  $K$  組のフォールト発見数データ  $(s_k, u_k, y_k)$  ( $k = 0, 1, 2, \dots, K$ ) が観測されたものとする．

はじめに，提案モデルから重回帰分析に必要な重回帰式を導出する．まず，式 (5) の両辺に対して自然対数をとることによって，

$$\log \gamma(s, u) = -\beta \log \rho + \alpha \beta \log s + (1 - \alpha) \beta \log u \quad (7)$$

を得る．これより，式 (7) は，次のような重回帰式に書き換えることができる．

$$Y_k = a_0 + a_1 K_k + a_2 L_k + \epsilon_k. \quad (8)$$

ここで，式 (8) において，

$$\begin{cases} Y_k = \log y_k, \\ K_k = \log s_k, \\ L_k = \log u_k, \\ a_0 = -\beta \log \rho, \\ a_1 = \alpha \beta, \\ a_2 = (1 - \alpha) \beta, \end{cases} \quad (9)$$

であり， $\epsilon_k$  は等分散性を仮定した実測データと回帰直線との残差を表す誤差項である．式 (8) から， $K$  組の実測データに対する誤差 2 乗和  $S(a_0, a_1, a_2)$  は，

$$\begin{aligned} S(a_0, a_1, a_2) &= \sum_{k=1}^K \epsilon_k^2 \\ &= \sum_{i=1}^K \{Y_k - (a_0 + a_1 K_k + a_2 L_k)\}^2 \end{aligned} \quad (10)$$

となる．したがって，式 (8) の  $a_0$  および回帰係数  $a_1$  および  $a_2$  の推定値  $\hat{a}_0, \hat{a}_1$ ，および  $\hat{a}_2$  は，それぞれ，式 (10) を最小にするときの値として推定される．すなわち， $\hat{a}_0, \hat{a}_1$ ，および  $\hat{a}_2$  は，同次連立方程式

$$\frac{\partial S}{\partial a_0} = \frac{\partial S}{\partial a_1} = \frac{\partial S}{\partial a_2} = 0 \quad (11)$$

を満足する値として得られる．以降，これらの方程式を数値的に解くことによって，推定値  $\hat{a}_0, \hat{a}_1$ ，および  $\hat{a}_2$  が得られ，最終的に，式 (5) に含まれるパラメータ  $\alpha, \beta$ ，および  $\rho$  の推定値  $\hat{\alpha}, \hat{\beta}$ ，および  $\hat{\rho}$  は，式 (9) から

$$\begin{cases} \hat{\alpha} = \frac{\hat{a}_1}{\hat{a}_1 + \hat{a}_2}, \\ \hat{\beta} = \hat{a}_1 + \hat{a}_2, \\ \hat{\rho} = \exp \left[ -\frac{\hat{a}_0}{\hat{a}_1 + \hat{a}_2} \right], \end{cases} \quad (12)$$

のように推定される．

ちなみに，式 (2) の従来の 1 変量ワイブル型モデルのパラメータは，単重回帰分析によって推定できる．まず，式 (2) の両辺に対して自然対数をそれぞれとることによって，次の式を得る．

$$\log \gamma(s) = -\beta \log \rho + \beta \log s. \quad (13)$$

式 (13) は、以下のような回帰式に書き換えることができる。

$$Y_k = b_0 + b_1 X_k + \epsilon_k. \tag{14}$$

ここで、式 (14) において、

$$\begin{cases} Y_k = \log y_k, \\ X_k = \log s_k, \\ b_0 = -\beta \log \rho, \\ b_1 = \beta, \end{cases} \tag{15}$$

である。したがって、式 (14) から、 $K$  組の実測データ  $(s_k, y_k)$  ( $k = 0, 1, 2, \dots, K$ ) に対する誤差 2 乗和  $S(a_0, a_1)$  を最小にする値  $\hat{b}_0$  および  $\hat{b}_1$  を求めた後、式 (15) の関係から、最終的に求めたいパラメータ推定値  $\hat{\beta}$  および  $\hat{\rho}$  を得ることができる。

### 6. 適用例

実際のテスト工程において観測されたフォールト発見数データを用いて、本論文において議論した 2 変量ワイブル型モデルに基づいたソフトウェア信頼性解析例を示す。本論文において適用する 2 つの実測データは、テスト時間の測定単位が週（カレンダー時間）であり、達成されたテスト網羅度（C0 カバレッジメジャー）に関するデータとともに記録された 24 組からなるフォールト発見数データ  $(s_k, u_k, y_k)$  ( $k = 1, 2, \dots, 24; s_{24} = 24$  (週),  $u_{24} = 0.9095, y_{24} = 296$ ) および同じく 22 組からなるフォールト発見数データ  $(s_k, u_k, y_k)$  ( $k = 1, 2, \dots, 22; s_{22} = 22$  (週),  $u_{22} = 0.9198, y_{22} = 212$ )<sup>20)</sup> である。本論文では、これらの 2 つの実測データを、それぞれ、DS1 および DS2 と呼ぶことにする。上記 2 つの実測データは、開発経験年数が 7 から 10 年目程度のソフトウェア開発技術者らが開発した組み込みソフトウェアのテスト工程で得られたフォールト発見数データであり、当該組み込みソフトウェアの開発規模は、それぞれ、 $1.972 \times 10^5$  (LOC: Lines of Codes) および  $1.630 \times 10^5$  (LOC) である。本論文では、テスト時間要因をカレンダー時間（週）、また、テスト労力要因には達成されたテスト網羅度を用いて提案モデルの適用例を示す。

はじめに、5 章において議論した提案モデルのパラメータ推定方法に基づいて、提案モデルに含まれるパラメータ  $\alpha, \beta$ , および  $\rho$  を推定する。まず、DS1 を用いて、式 (8) の  $a_0$  および回帰係数  $a_1$  および  $a_2$  を重回帰分析により推定した結果、 $a_0, a_1$ , および  $a_2$  の推定値  $\hat{a}_0, \hat{a}_1$ , および  $\hat{a}_2$  をそれぞれ得る。これらの推定値から、式 (12) を用いて、パラメータ推定値  $\hat{\alpha} = 0.17555, \hat{\beta} = 0.95652$ , および  $\hat{\rho} = 0.00415$  をそれぞれ得る。また、DS2 についても、

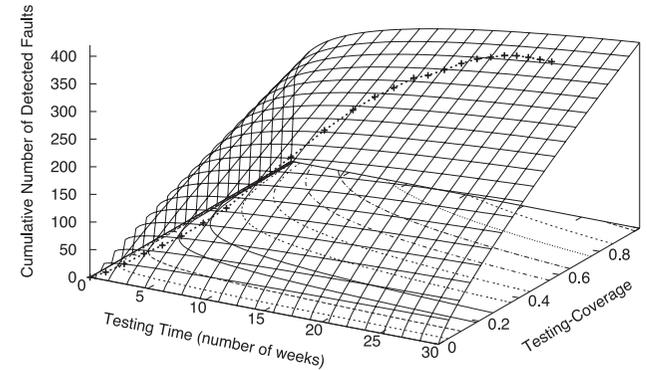


図 2 推定された 2 変量平均値関数 (DS1)  
Fig. 2 Estimated bivariate mean value function (DS1).

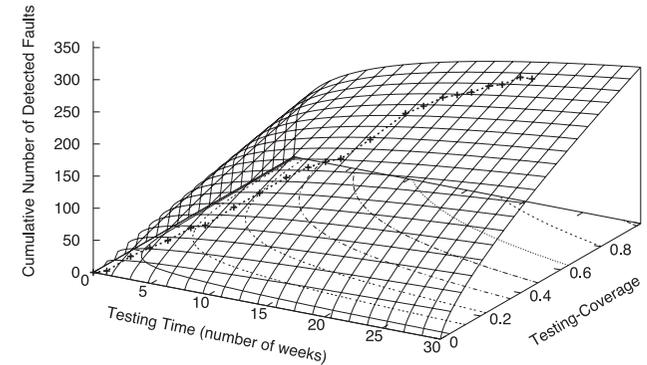


図 3 推定された 2 変量平均値関数 (DS2)  
Fig. 3 Estimated bivariate mean value function (DS2).

同様に、パラメータ推定値  $\hat{\alpha} = 0.29472, \hat{\beta} = 0.95336$ , および  $\hat{\rho} = 0.00856$  を得る。

図 2 および図 3 に、DS1 および DS2 を用いて推定された平均値関数を、それぞれ示す。ここで、2 次元時間空間に沿って描かれている 3 次元空間上の太破線は各時点において観測された実測データ、曲面は推定された 2 変量ワイブル型モデルの平均値関数  $\hat{\gamma}(s, u)$  の挙動を、それぞれ示している。また、2 次元時間空間上 (X-Y 平面上) に描かれている原点に対して凸な曲線は、推定された平均値関数の等高線を表しており、2 つの信頼度成長要因

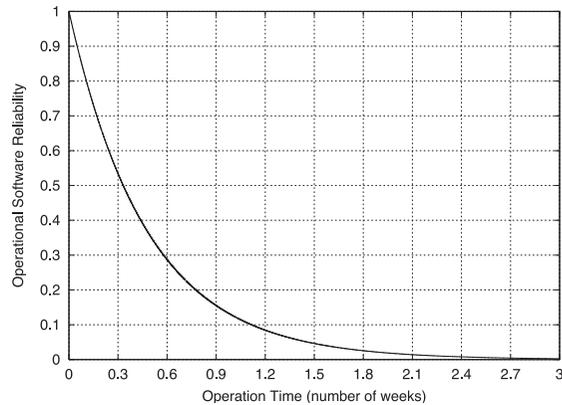


図4 推定されたソフトウェア運用信頼度  $\hat{R}(\eta | 24, 0.9095)$  (DS1)  
Fig. 4 Estimated operational software reliability,  $\hat{R}(\eta | 24, 0.9095)$  (DS1).

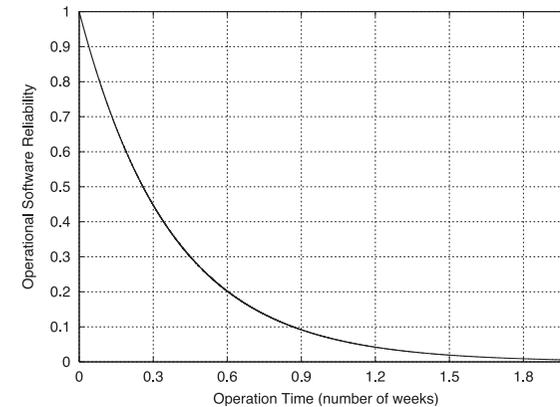


図5 推定されたソフトウェア運用信頼度  $\hat{R}(\eta | 22, 0.9198)$  (DS2)  
Fig. 5 Estimated operational software reliability,  $\hat{R}(\eta | 22, 0.9198)$  (DS2).

が当該ソフトウェア信頼度成長過程に及ぼす影響度を表現している。図2および図3から、今回提案したモデリング手法は、これまでの1変量SRGMとは異なり、テスト時間要因とテスト労力要因の2要因に依存しながらソフトウェア信頼度成長過程を曲面として描いていることが確認できる。また、推定された2次元平均値関数から、たとえ長い時間テストを実施したとしても、テスト網羅度が上昇しない限りソフトウェア故障は追加的に観測されにくく、除去されるフォールト数はほとんど増加しない状況や、一方、短いテスト時間において高いテスト網羅度を効率的に達成できた場合、それに応じて、ソフトウェア内に潜在的なフォールトが多く検出できるという状況など、1変量SRGMでは表現することができなかったより現実的なソフトウェア信頼度成長過程を表現していることが分かる。ところで、同じ等高線にある任意の2時点では、発見されるフォールト数が同じであるため、この2時点はお互いに無差別な時点として考えられる。したがって、上述のような平均値関数の等高線を、発見されたフォールト数に対する無差別曲線 (indifference curve for the number of detected faults) と呼ぶことができる。

次に、図4および図5に、推定されたソフトウェア運用信頼度  $\hat{R}(\eta | 24, 0.9095)$  および  $\hat{R}(\eta | 22, 0.9198)$  の時間的挙動をそれぞれ示す。たとえば、 $\hat{R}(\eta | 24, 0.9095)$  は、テスト開始後24週目までに達成されたテスト網羅度が0.9095であり、このときテストを終了してソフトウェアを出荷した場合、出荷時からのテスト時間区間  $(24, 24 + \eta]$  までにソフトウェ

ア故障が発生しない確率を表す。ただし、この場合、リリース後、テスト網羅度は出荷時のままで上昇せず、テスト時間要因のみがソフトウェア故障発生現象に影響を与えているというある意味強い仮定が存在していることに注意されたい。リリース後テスト網羅度が上昇しないという仮定は、運用段階よりもより厳しい環境を想定してテストを実施した場合に適用しうる仮定である。しかし、ソフトウェア品質保証の立場から一般的に考えれば、テストは運用段階よりも厳しい使用環境を想定して実施するべきであり、上述の仮定は、理論的にきわめて大きく乖離した仮定であるとはいえないものとする。図4および図5から、運用段階においてテスト終了時と同等の環境でソフトウェアを運用させた場合、テスト開始後1週目におけるソフトウェア運用信頼度  $\hat{R}(1.0 | 24, 0.9095)$  および  $\hat{R}(1.0 | 22, 0.9198)$  は、それぞれ、約0.12669および約0.07059と推定される。

## 7. モデルの適合性評価

6章で紹介した実測データを用いて、提案モデルの実測データに対する適合性および予測妥当性を総合的に検証する。はじめに、DS1およびDS2を用いて推定された式(8)の重回帰式について、適用した実測データに対する適合性をそれぞれ統計的に検証する。ここでは、DS1およびDS2を用いて推定された式(8)の重回帰式が、それぞれ、実測データをどの程度表現できているか統計的に分析するために、分散分析 (analysis of variance,

表 1 実測データに対する推定された重回帰式の分散分析表 (DS1)

Table 1 ANOVA table for the estimated multiple regression equation (DS1).

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$
Due to Regression	21.5091	2	10.7546	107,546**
Error	0.0021	21	0.0001	
Total	21.5112	23		

\*\* indicates significant of 1% level

表 2 実測データに対する推定された重回帰式の分散分析表 (DS2)

Table 2 ANOVA table for the estimated multiple regression equation (DS2).

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	$F_0$
Due to Regression	20.1505	2	10.0752	33,584**
Error	0.0059	19	0.0003	
Total	20.1564	21		

表 3 実測データに対する推定された重回帰式の重相関係数および寄与率

Table 3 Multiple coefficients of correlations ( $R$ ) and R-squares of the estimated multiple regression equations for the applied actual data.

	DS1	DS2
$R$	0.999952	0.99985
$R^2$	0.999904	0.99971
$R^*$	0.999947	0.99984
$R^{*2}$	0.999895	0.99967

以下 ANOVA と略す)<sup>21)–23)</sup> を用いてその検証を行う。ここで、DS1 および DS2 を用いて推定された重回帰式は、それぞれ、 $Y_i = 5.24617 + 0.167918K_i + 0.788604L_i$  および  $Y_i = 4.53912 + 0.28098K_i + 0.67238L_i$  である。表 1 および表 2 に、DS1 および DS2 を用いて推定された重回帰式の分散分析表をそれぞれ示す。表 1 および表 2 から、2 つの実測データを用いてそれぞれ推定された 2 つの重回帰式は、どちらとも実測データに対して高度に有意であることが分かる。すなわち、これらの推定された重回帰式は、それぞれ、実測データに対して適合しているといえる。

また、表 3 に、DS1 および DS2 を用いて推定された重回帰式と実測データとの重相関係数 (multiple coefficient of correlation:  $R$ ) および寄与率 (rate of contribution:  $R^2$ ) をそれぞれ示す。また、これらの適合度指標は、説明変数の自由度に依存するため、これらの

影響を除いた自由度調整済み重相関係数 (Adjusted R:  $R^*$ ) および自由度調整済み寄与率 (Adjusted R square:  $R^{*2}$ ) についても、それらの値を表 3 に示している。表 1, 表 2, および表 3 から、上述の推定された重回帰式は、それぞれの実測データに対して、統計的に非常に高い精度で適合していることがいえる。

さらに、2 章で紹介した 1 変量ワイブル型モデル (ONE-D WEIBULL), テスト網羅度を考慮した 1 次元 SRGM であるテスト網羅度依存型 SRGM (TCD)<sup>24)</sup>, 従来から多く実用に供されている NHPP モデルとして知られている指数形 SRGM<sup>10)</sup> (EXPO) および遅延 S 字形 (DELAYED-S)<sup>11)</sup> を取り上げ、平均偏差平方和 (mean squared error, 以下 MSE と略す) および予測相対誤差 (predicted relative errors, 以下 PRE と略す)<sup>3),4)</sup> の観点から提案モデルの実測データに対する適合性比較を行う。以下に、1 変量ワイブル型モデルを除いた上記 3 つの SRGM を示す。

- テスト網羅度依存型 SRGM :

$$\left. \begin{aligned} H(t) &= a [1 - \exp \{-r \cdot C(t)\}] \\ C(t) &= \frac{\alpha(1 - e^{-bt})}{1 + z \cdot e^{-bt}} \quad (a > 0, r > 0, b > 0, z > 0, 0 < \alpha \leq 1.0). \end{aligned} \right\} \quad (16)$$

- 指数形 SRGM :

$$H(t) = a(1 - e^{-bt}) \quad (a > 0, b > 0). \quad (17)$$

- 遅延 S 字形 SRGM :

$$H(t) = a \{1 - (1 + bt)e^{-bt}\} \quad (a > 0, b > 0). \quad (18)$$

ここで、 $a$  はテスト開始前にソフトウェア内に潜在する総期待フォールト数、 $r$  は任意のテスト時刻  $t$  において向上した単位テスト網羅度あたりのフォールト検出率、 $z$  はテストケース設計者のテスト習熟性を示す能力係数、および  $b$  はフォールト 1 個あたりのフォールト発見率を、それぞれ表す。

MSE は、 $K$  組のフォールト発見数データが観測されている場合、ある一定のテスト時刻  $\Psi_k$  ( $k = 1, 2, \dots, K$ ) までに発見された総フォールト数の実測値  $y_k$  と推定値  $\hat{y}(\Psi_k)$  との偏差 2 乗和をデータ数で平均化したものであり、

$$\text{MSE} = \frac{1}{K} \sum_{k=1}^K \{y_k - \hat{y}(\Psi_k)\}^2 \quad (19)$$

表 4 MSE に基づくモデルの適合性比較結果  
Table 4 Results of model comparisons based on MSE.

	DS1	DS2
OUR MODEL	<b>1.96356</b>	<b>3.09947</b>
ONE-D WEIBULL	1,425.24*	205.688*
TCD	111.838	49.9692
EXPO	603.198	73.0885
DELAYED-S	150.005	36.6406

として算出される。したがって、小さな MSE 値を示すモデルほど、適用した実測データに対して適合性が高いモデルであると判断される。ここで、 $\Psi_k$  は、時間空間ベクトルの要素を表している。したがって、1 次元時間空間の場合は  $\Psi_k \equiv s_k$ 、2 次元時間空間の場合は  $\Psi_k \equiv s_k, u_k$  となる。また、PRE は、任意のテスト時刻  $\Psi_q$  ( $0 \leq \Psi_q \leq \Psi_e$ ) までの観測データを用いて推定したときの、テスト終了時刻  $\Psi_e$  までに発見される総フォールト数の予測値  $\hat{y}(\Psi_q; \Psi_e)$  と実測値  $y_e$  との相対誤差を示す。ここで、任意のテスト時刻  $\Psi_q$  における PRE を  $PRE[\Psi_q]$  とすると、

$$PRE[\Psi_q] = \frac{\hat{y}(\Psi_q; \Psi_e) - y_e}{y_e} \quad (20)$$

として計算される。したがって、テスト早期段階から安定的にゼロに近い PRE 値を示すモデルほど、予測性能が優れたモデルであると判断される。

表 4 および表 5 に、MSE および PRE に基づいたモデルの適合性比較結果をそれぞれ示す。なお、表 4 および表 5 における『\*』は、パラメータ推定の結果、理論的に適切なパラメータ推定値が得られず、MSE 値もしくは PRE 値が算出できたとしても評価に値しないことを意味する。『—』は、パラメータ推定の結果、推定値が得られず MSE 値および PRE 値が計算できないことを示す。また、TCD、EXPO、および DELAYED-S の各モデルのパラメータは、最尤法を用いて推定している。ところで、表 5 では、DS1 および DS2 に対して、それぞれ、テスト進捗率 (Testing-progress ratio) 50%時点から 5%の間隔で PRE 値を計算した結果を示している。これは、実際のテスト工程における信頼性評価開始時期を考慮したためであり、通常、実測データの挙動を見極めるため、テスト進捗率 60~70%時点以降から SRGM に基づく信頼性評価を開始することが妥当だとされている。

まず、表 4 から、提案モデルは、本論文において取り上げた 2 つの実測データがそれぞれ示すソフトウェア信頼度成長過程を非常に高い精度で表現できていることが分かる。また、表 5 では、提案モデルは、本論文において取り上げた既存の 1 変量 SRGM と比較し

表 5 PRE に基づくモデルの適合性比較結果  
Table 5 Results of model comparisons based on PRE.

DS1	OUR MODEL	ONE-D WEIBULL	TCD	EXPO	DELAYED-S
TPR (%)					
50	<b>2.311E-02</b>	7.812E-01*	-1.052E-01*	4.970E-01	3.951E-01
55	<b>2.544E-02</b>	7.803E-01*	-1.201E-01*	4.620E-01	2.619E-01
60	<b>2.838E-02</b>	7.628E-01*	-9.306E-02*	4.336E-01	2.029E-01
65	<b>2.958E-02</b>	7.271E-01*	-1.006E-01*	3.678E-01	1.108E-01
70	<b>3.034E-02</b>	6.855E-01*	-7.841E-02*	3.295E-01	8.602E-02
75	3.148E-02	5.999E-01*	<b>-2.699E-02</b>	2.150E-01	7.618E-02
80	3.035E-02	5.555E-01*	<b>-1.602E-03</b>	1.568E-01	5.957E-02
85	2.887E-02	5.116E-01*	<b>-3.459E-03</b>	1.183E-01	5.108E-02
90	2.702E-02	4.683E-01*	<b>2.898E-03</b>	8.231E-02	3.902E-02
95	2.384E-02	4.683E-01*	<b>3.379E-03</b>	4.867E-02	2.406E-02
100	1.476E-02	3.423E-01*	0	1.920E-16	0
DS2	OUR MODEL	ONE-D WEIBULL	TCD	EXPO	DELAYED-S
TPR (%)					
50	<b>2.477E-02</b>	5.384E-01*	-3.554E-01	8.743E-02	-6.164E-01
55	3.441E-02	4.587E-01*	-3.720E-01	<b>3.053E-03</b>	-5.332E-01
60	<b>3.795E-02</b>	4.139E-01*	-1.830E-01	9.600E-02	-4.875E-01
65	<b>3.570E-02</b>	3.940E-01*	—	1.887E-01	-4.534E-01
70	<b>3.161E-02</b>	3.804E-01*	8.070E-02*	1.717E-01	-3.504E-01
75	<b>2.597E-02</b>	3.604E-01*	3.409E-02*	1.629E-01	-2.670E-01
80	<b>2.099E-02</b>	3.350E-01*	-2.238E-02*	1.264E-01	-1.997E-01
85	<b>1.683E-02</b>	3.071E-01*	-3.695E-02	8.779E-02	-1.492E-01
90	<b>1.260E-03</b>	2.811E-01*	-1.696E-02	7.055E-02	-9.719E-02
95	<b>9.523E-03</b>	2.541E-01*	-1.980E-02	3.312E-02	-6.373E-02
100	5.137E-03	2.048E-01*	-6.880E-05	-2.434E-07	3.545E-07

(TPR: Testing-progress ratio, E: Exponential in decimal)

て、通常の信頼性評価開始時点よりも早い段階から、テスト終了時刻までに発見される総期待フォールト数を安定的かつ精度良く予測できていることが分かる。次に、表 4 および表 5 において得られた上述した適合性比較結果について、いくつかの考察を与える。まず、図 6 および図 7 に、本論文において取り上げた 1 変量 SRGM について、DS1 および DS2 を用いて推定された総期待発見フォールト数の挙動をそれぞれ示す。これらの図より、適用する SRGM によって総期待発見フォールト数の推定値の時間的挙動は大きく異なることが分かる。具体的に、図 6 では、1 次元時間空間に沿って示された DS1 は明らかに S 字形信頼度成長曲線を示しており、元々 S 字形の性質を有しない指数形 SRGM や 1 変量ワイブル型モデルの推定精度は悪いことが分かる。また、DS2 は、DS1 とは多少異なった S 字形信頼度

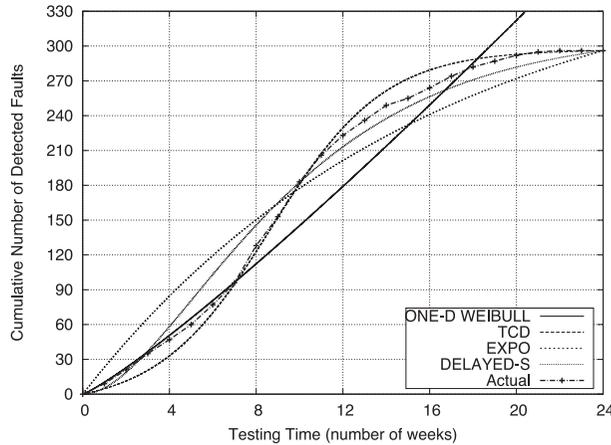


図 6 推定された 1 変量 SRGM の平均値関数 (DS1)

Fig. 6 Estimated mean value functions of the one-dimensional SRGMs (DS1).

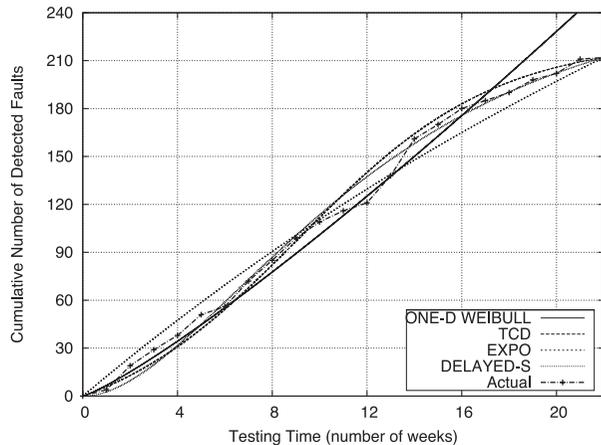


図 7 推定された 1 変量 SRGM の平均値関数 (DS2)

Fig. 7 Estimated mean value functions of the one-dimensional SRGMs (DS2).

成長曲線を示す実測データであり、テスト開始後 12 週目あたりにチェンジポイント<sup>25)</sup>が発生している様子がうかがえる。DS2 を用いた場合には、実測データが示す信頼度成長曲

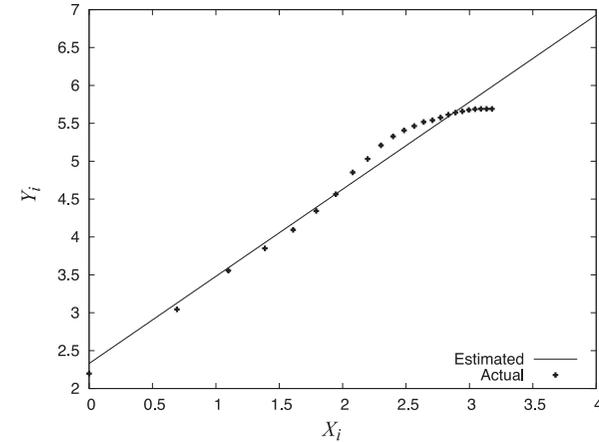


図 8 1 変量ワイブル型モデルに基づく推定された回帰式 (DS1)

Fig. 8 Estimated regression equation derived from the one-dimensional Weibull-type model (DS1).

線の形状やチェンジポイントなどが MSE や PRE に基づく適合性比較結果に少なからず影響を与えているものと考えられる。しかしながら、1 次元時間空間上の実測データが示す信頼度成長曲線の特徴が SRGM の推定精度に与える影響は、2 次元時間空間上では実測データ自身の挙動を説明する要因が増えるため、それはほとんど緩和される。ところで、MSE および PRE による適合性比較において特に結果が悪くなかった 1 変量ワイブル型モデルは、図 6 および図 7 を見ると、テスト早期段階を除き、実測データの挙動をほとんどうまく表現できていないことが分かる。これは、パラメータ推定の際に導出した回帰式が、実測データをどれだけ精度良く表現できているかに依存している。図 8 および図 9 に、それぞれ、DS1 および DS2 を用いて推定された式 (14) の回帰式を示す。これより、1 変量ワイブル型モデルに基づく推定された回帰式は、テスト早期段階では実測データが示す挙動をある程度表現できているものの、中・後期段階ではほとんどうまく表現できていないことが分かる。この状況は、推定された 1 変量ワイブル型モデルの平均値関数の振舞いに影響を与えていることが、図 6 および図 7 から確認できる。また、DS1 において、テスト網羅度依存型 SRGM は、1 変量 SRGM の中でも MSE 値およびテスト後期段階における PRE 値は良い値を示しているが、これは説明変数の数が他の 1 変量モデルに比べてきわめて多いことが原因として考えられる。ただし、本論文における適合性比較結果から総合的に分かるように、理論的に適切なパラメータ推定値を同時に得られないことやテスト早期段階において安

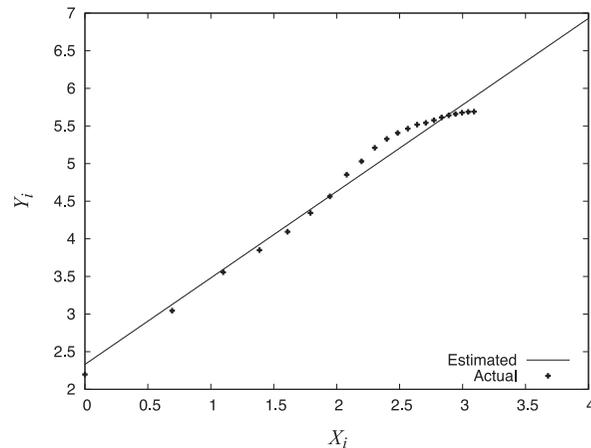


図9 1変量ワイブル型モデルに基づく推定された回帰式 (DS2)

Fig.9 Estimated regression equation derived from the one-dimensional Weibull-type model (DS2).

定的な PRE 値が得られないなど、その弊害も現れている。一方、表1、表2、および表3から分かるように、DS1 および DS2 を用いて推定された提案モデルに基づく重回帰式は実測データを高精度に推定できており、モデルに含まれるパラメータ数が少ないにもかかわらず、高精度な信頼性評価を実現している。

## 8. おわりに

本論文では、コブ・ダグラス型効用関数の考え方に基づいて、従来のワイブル型モデルにおけるソフトウェア信頼度成長要因としての1次元時間空間をテスト時間要因とテスト労力要因の2要因からなる2次元時間空間へ拡張した2変量ワイブル型モデルの構築を行った。さらに、組み込みソフトウェアのテスト工程で得られた2つの実測データを用いながら、提案モデルに基づくソフトウェア信頼性解析例を示した。特に、本論文では、提案モデルの実測データに対する適合性を統計的に分析するとともに、平均偏差平方和や予測相対誤差を評価基準として用いながら、提案モデルと既存の代表的な1変量SRGMとの実測データに対する適合性比較を行い、少なくとも今回適用した実測データに対しては、提案モデルを用いることで高精度のソフトウェア信頼性評価が行えることを検証し、今回得られた適合性比較結果に関する考察を与えた。

従来の1変量ワイブル型モデルでは、テスト工程におけるソフトウェア信頼度成長過程をテスト時間のみに依存したものととらえてきたが、本論文で提案したような2変量ワイブル型モデルは、ソフトウェア信頼度成長過程が前述の2要因に依存して説明されるため、理論的にも現実的にも整合性を有したソフトウェア信頼性評価が可能となる。また、テスト時間要因とテスト労力要因からなるソフトウェア信頼度成長要因が信頼度成長過程に与える影響度合いをコブ・ダグラス型効用関数の概念を用いて表現することによって、その一般性を失うことなく様々なタイプの2次元ソフトウェア信頼度成長過程を記述できる利点を持つ。今後は、テスト網羅度以外のテスト労力要因データを用いた場合における提案モデルの信頼性評価性能を検証する必要があると同時に、2変量SRGMに基づいた定量的なソフトウェア信頼性評価に有用な信頼性評価尺度の新たな開発とソフトウェアの最適リリース問題などソフトウェア開発管理面からの種々の諸問題についても議論を行い、2変量ソフトウェア信頼性評価技術としてそれらを体系的に確立していく必要がある。

謝辞 本研究の一部は、日本学術振興会科学研究費補助金若手研究(B)(課題番号19710129)、基盤研究(C)(課題番号18510124)、および鳥取大学ベンチャー・ビジネス・ラボラトリー平成19年度提案型研究開発テーマ研究助成の援助を受けたことを付記する。

## 参考文献

- 1) Yamada, S. and Osaki, S.: Software reliability growth modeling: Models and applications, *IEEE Trans. Softw. Eng.*, Vol. SE-11, No.12, pp.1431-1437 (1985).
- 2) Musa, J.D., Iannio, D. and Okumoto, K.: *Software Reliability: Measurement, Prediction, Application*, McGraw-Hill, New York (1987).
- 3) 山田 茂: ソフトウェア信頼性モデル—基礎と応用, 日科技連出版社 (1994).
- 4) Pham, H.: *Software Reliability*, Springer-Verlag, Singapore (2000).
- 5) 山田 茂, 藤原隆次: ソフトウェアの信頼性: モデル, ツール, マネジメント, プロジェクトマネジメント学会 (PM学会教育・出版シリーズ(1)) (2004).
- 6) 山田 茂, 福島利彦: 品質指向ソフトウェアマネジメント, 森北出版 (2007).
- 7) Yamada, S., Ohtera, H. and Narihisa, H.: Software reliability growth models with testing-effort, *IEEE Trans. Reliab.*, Vol. R-35, No.1, pp.19-23 (1986).
- 8) Fujiwara, T. and Yamada, S.: Software reliability growth modeling based on testing-skill characteristics: Model and Application, *Elec. Commu. Japan (Part 3)*, Vol.84, No.6, pp.42-49 (2001).
- 9) Huang, C.Y.: Performance analysis of software reliability growth models with testing-effort and change-point, *J. Syst. Softw.*, Vol.76, No.2, pp.181-194 (2005).
- 10) Goel, A.L. and Okumoto, K.: Time-dependent error-detection rate model for soft-

- ware reliability and other performance measures, *IEEE Trans. Reliab.*, Vol.R-28, No.3, pp.206–211 (1979).
- 11) Yamada, S., Ohba, M. and Osaki, S.: S-shaped reliability growth modeling for software error detection, *IEEE Trans. Reliab.*, Vol.R-32, No.5, pp.475–478, 485 (1983).
  - 12) 石井智隆, 土肥 正: 二次元 NHPP に基づいたテスト労力依存型ソフトウェア信頼性モデル, 電子情報通信学会論文誌, Vol.J89-D, No.8, pp.1684–1694 (2006).
  - 13) 井上真二, 山田 茂: ワイブル過程に基づいた2次元ソフトウェア信頼度成長モデル, 電子情報通信学会技術研究報告, Vol.107, No.137, pp.39–44 (2007).
  - 14) Keiller, P.A. and Miller, D.R.: On the use and the performance of software reliability growth models, *Rel. Eng. Syst. Safe.*, Vol.32, No.1-2, pp.95–117 (1991).
  - 15) Varian, H.R.: *Intermediate Microeconomics – A Modern Approach, 2nd Edition*, W.W. Norton & Company, New York (1991).
  - 16) Ahn, C.W., Chae, K.C. and Clark, G.M.: Estimating parameters of the power law process with two measures of failure time, *J. Qual. Tech.*, Vol.30, No.2, pp.127–132 (1998).
  - 17) Murthy, D.N.P., Iskandar, B.P. and Wilson, R.J.: Two-dimensional failure-free warranty policies: Two-dimensional point process models, *Opera. Res.*, Vol.43, No.2, pp.356–366 (1995).
  - 18) Murthy, D.N.P., Baik, J., Wilson, R.J. and Bulmer, M.R.: Two-dimensional failure modeling, in *Springer Handbook of Engineering Statistics* (Pham, H. Ed.), pp.97–111, Springer-Verlag, Berlin (2006).
  - 19) Yamada, S., Hishitani, J. and Osaki, S.: Software-reliability growth with a Weibull test-effort, *IEEE Trans. Reliab.*, Vol.42, No.1, pp.100–106 (1993).
  - 20) Fujiwara, T. and Yamada, S.: C0 coverage-measure and testing-domain metrics based on a software reliability growth model, *Intern. J. Rel. Quali. Safe. Eng.*, Vol.9, No.4, pp.329–340 (2002).
  - 21) 山田 茂, 木村光宏, 高橋宗雄: TQM のための統計的品質管理—一般工業製品からソフトウェア品質まで, コロナ社 (1998).
  - 22) 永田 靖, 棟近雅彦: 多変量解析入門, サイエンス社 (2001).
  - 23) Montgomery, D.C.: *Introduction to Statistical Quality Control*, John Wiley & Sons, New Jersey (2005).
  - 24) Inoue, S. and Yamada, S.: Testing-coverage dependent software reliability growth modeling, *Intern. J. Rel. Quali. Safe. Eng.*, Vol.11, No.4, pp.303–312 (2004).
  - 25) Inoue, S. and Yamada, S.: Software reliability measurement with change-point, *Proc. 5th Intern. Conf. Quali. Reliab.*, pp.170–175 (2007).

(平成 19 年 10 月 15 日受付)

(平成 20 年 5 月 8 日採録)



井上 真二 (正会員)

平成 13 年鳥取大学工学部社会開発システム工学科卒業。平成 18 年同大学大学院工学研究科博士後期課程修了。平成 18 年鳥取大学工学部社会開発システム工学科助手。平成 19 年同大学工学部社会開発システム工学科助教。平成 20 年同大学工学研究科社会基盤工学専攻助教, 現在に至る。博士(工学)。ソフトウェア信頼性工学, プロジェクトマネジメントに関する研究に従事。2004 International Computer Symposium – Best Paper Award, IEEE Reliability Society Japan Chapter – 2006 Outstanding Young Scientist Award 受賞。日本オペレーションズ・リサーチ学会, 電子情報通信学会, 日本品質管理学会, プロジェクトマネジメント学会, IEEE 各会員。



山田 茂 (正会員)

昭和 50 年広島大学工学部経営工学科卒業。昭和 52 年同大学大学院工学研究科修士課程修了。昭和 52~54 年日本電装(株)品質保証部勤務。昭和 58 年広島大学大学院博士課程修了。昭和 58~63 年岡山理科大学勤務。昭和 63 年広島大学工学部第二類助教授。平成 5 年鳥取大学工学部社会開発システム工学科教授。平成 20 年同大学大学院工学研究科社会基盤工学専攻教授, 現在に至る。工学博士。ソフトウェア信頼性工学, 品質管理工学, プロジェクトマネジメントの研究に従事。情報処理学会平成 3 年度 Best Author 賞, 第 8 回電気通信普及財団賞(テレコムシステム技術賞), 日本信頼性学会平成 10 年高木賞, International Readership Award in Reliability Engg. Research (ICQRIT/SRECOM, 2003), 2004 International Computer Symposium – Best Paper Award, 平成 17 年度プロジェクトマネジメント学会論文賞, Leadership Award (ISSAT, 2007) 等受賞。著書『ソフトウェアマネジメントモデル入門』(共立出版, 1993), 『ソフトウェア信頼性モデル—基礎と応用』(日科技連出版社, 1994), 『TQM のための統計的品質管理』(コロナ社, 1998), 『ソフトウェア信頼性: モデル, ツール, マネジメント』(プロジェクトマネジメント学会, 2004), 『品質指向ソフトウェアマネジメント』(森北出版, 2007) ほか。日本オペレーションズ・リサーチ学会(フェロー), 電子情報通信学会, 日本応用数理学会, 日本経営工学会, 日本信頼性学会, 日本品質管理学会, プロジェクトマネジメント学会, IEEE 各会員。