

途中工程データが欠落した状況における影響評価モデルの構築

## Establishment of influence evaluation model under the circumstances

### that defect data is not available in some development phases

キヤノン株式会社 医療機器システム評価室  
Canon Inc. Medical Equipment System Assurance Dept.  
加藤 拓海  
Takumi Kato

#### Abstract

In software development, it is well known that upstreaming the detection of defects improves the quality in both the downstream and the field. It is often assumed that the number of defects in testing process is proportional to the size of code. However, upstream activities including reviews must have some influence on the number of defects in testing phases. In this report, using actual project data, we build an influence evaluation model by making use of a front-loading ratio of defect detection. It quantitatively shows how such activities in upstream gives an impact on quality in test phases. The contribution of this paper defines establishment of influence evaluation model under the circumstances that defect data is not available in some development phases.

#### 1. はじめに

一般にソフトウェア開発においてより上流で欠陥を検出することは、下流のテスト工程での品質を向上し、手戻りを減らし、市場への流出低減にも効果があることが知られている[1]。筆者の所属する組織（以下、当組織）においても上流品質改善の取り組みが活発となり、レビューで抽出された欠陥をすべて記録に残す運用が定着し、レビューの質を向上する施策も行われてきた。感覚的には品質が向上したものの、上流工程が下流工程に与える影響が定量的に示されていないため、その効果の度合いを計れずにいた。また、この定量化の過程において、単にレビューで抽出された欠陥の数（以降、レビュー欠陥数とする）で目標設定してしまうと、欠陥の件数に意識が向いて、下流工程への流出防止効果が十分に発揮されない懸念がある。このため、どれだけ上流で欠陥を検出できたかを示す指標を定義し、これを用いた目標設定が必要と考えられる。

先行研究[2][3]では、ある工程でのレビューがどれだけ徹底して行われたかを示す指標として前倒し率を定義し、これを用いて、上流でのレビュー実績が下流でのテスト欠陥数に与える効果を定量的に表現したモデルを構築している。更に、前倒し率を近似変換することによって、上流工程終了時に下流工程での品質を予測可能な品質予測モデルへと進化させている。

本研究では、この2つのモデルを当組織におけるプロジェクトへ適用可能であるかを検証し、定量的モデルを構築する。これを影響評価モデルと呼ぶことにする。適用にあたり、当組織におけるプロジェクトでは途中工程のデータが収集できていない状況があり、先行研究のやり方をそのままでは適用できない。工程データが収集できない1つの理由は、欠陥数が多く記録が大変なためである。このような状況は実際によく起きるため、定量的モデルを構築できれば、実用上の価値は高いと考えられる。そこで、本論文での研究課題を次の通り設定する。

【研究課題】 途中工程データが欠落している状況における影響評価モデルの構築

---

キヤノン株式会社 医療機器システム評価室  
Medical Equipment System Assurance Dept., Canon Inc.

神奈川県川崎市中原区今井上町 9-1 Tel: 03-5758-2111 e-mail:kato.takumi@canon.co.jp  
9-1, Imai-kamicho, Nakahara, Kawasaki, Kanagawa Japan

## 2. 現状の課題

### 2.1 目指すべき姿

本研究の対象は、メトリクスの収集を開始してから現在までに完了している6プロジェクトである。ウォーターフォール型の開発を行っており、上流から、要件検討/仕様検討/設計検討/実装/単体テスト/結合テスト/機能テスト/システムテストという工程を採用している。このうち要件～設計フェーズはドキュメントを対象としたレビューが必須化されており、工数や欠陥数などが計測されている。また結合テスト以降のフェーズではすべての欠陥が計測されている。

ここで2つのプロジェクトにおける各工程での欠陥割合を図1、2に示す。実装と単体工程が抜けているのは、該当工程における欠陥の記録が義務付けられていないためである。2つのグラフを比較すると、上流で多く欠陥を除去できたプロジェクトは、下流工程の品質が向上していることが見て取れる。Prj. Fでは上流3工程で抽出した割合（以下、上流抽出率）が54%に留まり、結合テストで31%も抽出するも、システムテストへ6%流出させている。一方、Prj. Eでは上流で83%を除去し、システムテストへの流出を1%に抑え込んでいる。NECの品質会計においては、上流工程での欠陥抽出が80%となることを目標としている[1]。以上から、Prj. Eのような状況を理想的な姿として、これに近づけるように開発プロセスの各過程において適切なアクションを選択できるようにしたい。

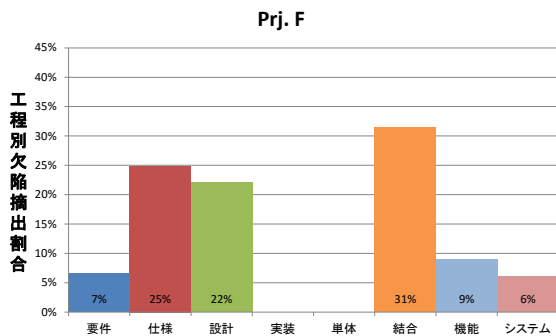


図1：各工程における欠陥抽出割合①

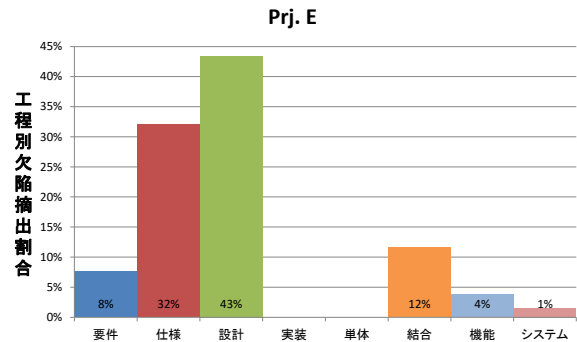


図2：各工程における欠陥抽出割合②

### 2.2 変更規模で予測することの課題

テスト工程での欠陥数が変更規模に比例することは一般に広く知られている[4]。当組織においてもこれまで欠陥密度、つまり変更規模当たりの欠陥数を用いて予測を行ってきた。対象プロジェクトの変更規模とテスト欠陥数に対して線形回帰分析した結果、当て嵌まりのよさを表す自由度調整済み決定係数は68%と物足りない。散布図（図5）で確認したところ1点大きく外れた点があった。状況を確認したところ、上流での欠陥抽出が十分でないため、テスト工程で大規模な施策を実施していることがわかった（図6）。このように、上流と下流の欠陥検出の比が通常と大きく異なる場合には単に変更規模との比例関係だけでは説明できない課題があることがわかった。

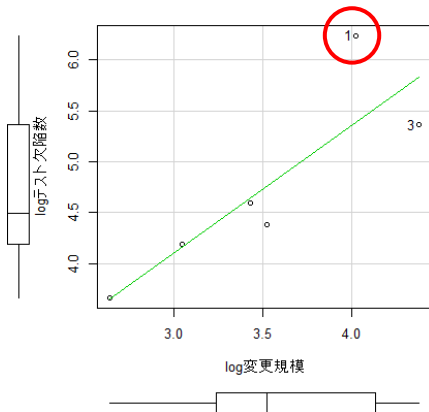


図5：散布図

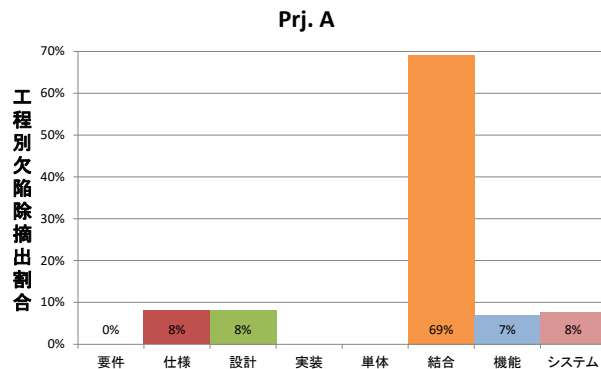


図6：外れ値プロジェクトにおける欠陥抽出割合

## 2.3 実現したいこと

### (1) 上流工程終了時におけるアクションの検討

上流でのレビュー活動終了時点で、レビュー実績から下流品質を予測し、目標に達していない場合は必要なアクションを取れるようにしたい。イメージを図3に示す。まず統計解析によりレビュー欠陥数と上流摘出率の関係をグラフ化する。これを用いて、上流工程終了時のレビュー欠陥数の実績値から、欠陥検出における上流摘出率を予測する。目標値未達の場合は、追加レビューであと何件の欠陥を摘出すればよいか分かる。テスト工程での欠陥数(以下、テスト欠陥数)ではなく上流摘出率で示すことで、プロセス全体を意識した議論を促す。

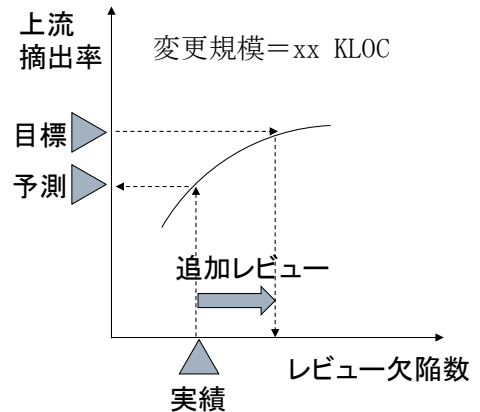


図3：実現したいこと(1)イメージ

### (2) プロジェクト開始時におけるレビュー目標の設定

プロジェクトの開始時点で、下流工程への影響を踏まえたレビュー欠陥数の目標設定を実現したい。イメージを図4に示す。まず統計解析により、上流摘出率を固定した際の変更規模とレビュー欠陥数の関係をグラフ化する。これを用いて、プロジェクト開始時の変更規模の見積り値から、目標とする上流摘出率を達成するためにレビューで摘出すべき欠陥数を得る。(1)同様、上流摘出率を示すことで、下流工程を意識したレビュー活動を促す。

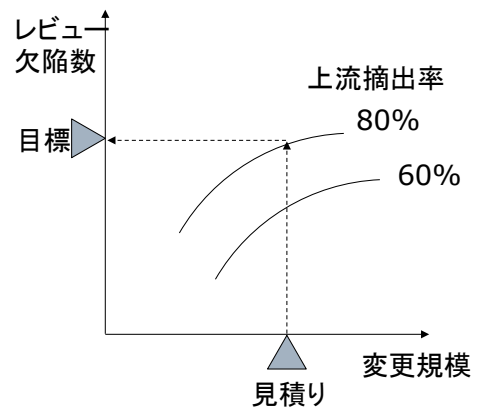


図4：実現したいこと(2)イメージ

## 3. 影響評価モデルの構築

### 3.1 先行研究

先行研究[2]では、上流工程のレビュー欠陥数から下流工程でのテスト欠陥数を予測する品質予測モデルが提案されている。ある工程でのレビューの徹底度合いとして、その工程での欠陥摘出数をそれ以降の工程での欠陥摘出数との比で表した前倒し率という指標を定義している。ドキュメントレビューを対象とした前倒し率を図7に示す。

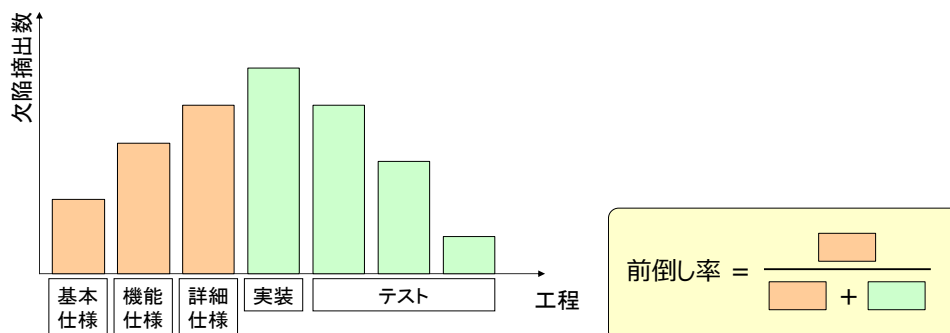


図7：先行研究における前倒し率の定義

前倒し率は、テスト工程の欠陥数に指数関数を通して影響することが、ワイブル分布（特にレイリー分布）の性質から示されている。工程別の欠陥数に影響する要因としては、開発規模、難易度、類似システムの経験度合い、そしてレビューによる品質向上効果などが検討されている。そのうち最初の3要因は各工程の欠陥数に対して同じ方向（増大／減少）に作用するのに対して、レビューの品質向上効果は上流では増大、下流では減少する向きに作用するため、前倒し率を用いることでレビューの品質向上効果のみをうまくとりだせることが論じられている。

モデル構築の過程では、まず①効果説明モデルを構築している。これは、前倒し率を高めれば下流のテスト工程での欠陥検出数が、開発規模などから予想される値より低下するという関係を定量的に表現したモデルである。ただし前倒し率の定義から明らかなように、その算出には目的変数でもあるテスト欠陥数が必要であるため、プロジェクトが終了するまで確定しない。よって次に、①を基に前倒し率を近似変形して、テスト欠陥数を入力としない②品質予測モデルを構築している。

これらに対して、変数名を適宜、本研究での呼び方に置き換えたものを以下に示す。ただし、[2]ではプロジェクトリスク情報としていくつかのダミー変数を説明変数に加えているが、ここでは省略している。

$$\log(\text{テスト欠陥数}) = \alpha_1 \log(\text{変更規模}) + \alpha_2 \text{前倒し率} + c \quad \textcircled{1}$$

$$\log(\text{テスト欠陥数}) = \alpha_1' \log(\text{変更規模}) + \alpha_2' \frac{\text{変更規模}}{\text{レビュー欠陥数}} + c \quad \textcircled{2}$$

本研究ではこの2つのモデルについて、対象プロジェクトへの適用が可能であるかを検証する。

### 3.2 本研究における前倒し率と研究仮説

先行研究に倣い、本研究における前倒し率を定義する。ただし、記録のない工程があるため、先行研究と同じ定義は使えない。まず上流の3工程はいずれもドキュメントを対象としたレビューを行うもので、そのまま適用できる。実装工程と単体テスト工程におけるデータが欠落しているため、本研究における前倒し率の定義は図8のようになる。

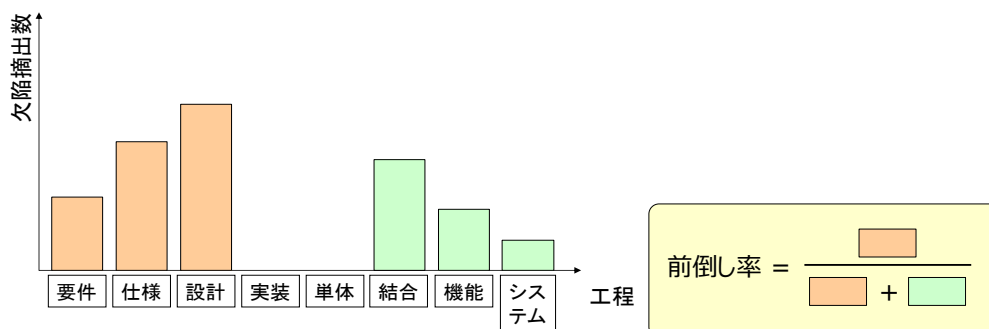


図8：本研究における前倒し率の定義

このように、先行研究とは指標の定義が異なるが、工程別の欠陥数の比をとっている点では共通している。特に、先行研究での議論と同様に、開発規模、難易度、経験度などの他の要因の影響を受けにくく、レビューの品質向上効果をあらわす指標になっていることは期待できる。加えて②品質予測モデルでは、開発規模がわかれば、実装および単体工程での欠陥数はモデルの入力として不要であり、我々の状況でも同じように成立している可能性がある。

以上を研究仮説としてまとめると次のようになる。

(仮説A) ①の前倒し率を本研究の定義に置き換えても定量的モデルが確立できる

(仮説B) ②の形のモデルは本研究の状況でも確立できる

### 3.3 プロジェクトデータを用いた検証

3.1 節で述べた先行研究における 2 つのモデル①②に対して、対象プロジェクトでの実績データを用いて回帰分析を行った。先行研究に従い、ポアソン分布を用い、当て嵌まりの良さを表す指標は McFadden の疑似  $R^2$  としている。結果をそれぞれ表 1 の①②に示す。

表 1：前倒し率を用いた分析結果

	偏回帰係数	係数 p 値	疑似 $R^2$
モデル①	-2.1	0.0015	0.99
モデル②	2.8	<2e-16	0.99

まず、2.2 節で外れ値としたデータの考慮が不要となっている。これは変更規模による影響だけでは説明がつかず特異と見られたデータが、前倒し率の概念を導入したことにより説明が可能となったことを意味する。モデル①における偏回帰係数の p 値は 1%有意、疑似  $R^2$  は約 99%と申し分ない結果であった。対象プロジェクトでは途中工程のデータが欠落しているため、本研究における前倒し率の定義には懸念があった。しかし①の結果はそれを払拭し、途中工程データが欠落している場合においても適用可能であることを示しており、仮説 A の成立が確認された。

次にモデル②についても同様に、偏回帰係数の p 値は 1%有意、疑似  $R^2$  は約 99%であり、仮説 B の成立が確認された。

ここで、表 1 では有効桁数 2 桁としたため疑似  $R^2$  は同じ値となっているが、分析ツールの出力値で比較すると①に比べ②の方が 0.1%程よい結果となった。先行研究では近似変形を行う過程で精度が低下しているのに対し、本検証では僅かながら向上していることについて考察する。実装および単体工程がソースコードを直接扱うという意味で、欠落しているデータと変更規模との関連性は強いと考えられる。モデル②では前倒し率を変更規模とレビュー欠陥数との比に置き換えている。モデル①では欠落データを除外したが、変更規模をより活用するモデル②によってその影響が取り込まれ、結果として精度が向上したと捉えることができる。

以上の考察を踏まえ、本研究の成果としてモデル②を採用する。得られた係数を代入し、左辺に対数関数を含まない形に変形した以下の式を、本研究における影響評価モデルとする。

#### 【影響評価モデル】

$$\text{テスト欠陥数} = 1.7315 * \text{変更規模}^{1.01795} * \exp\left(2.78053 \frac{\text{変更規模}}{\text{レビュー欠陥数}}\right)$$

この予測式によるテスト欠陥数予測と実際のテスト欠陥数の散布図を図 9 に示す。

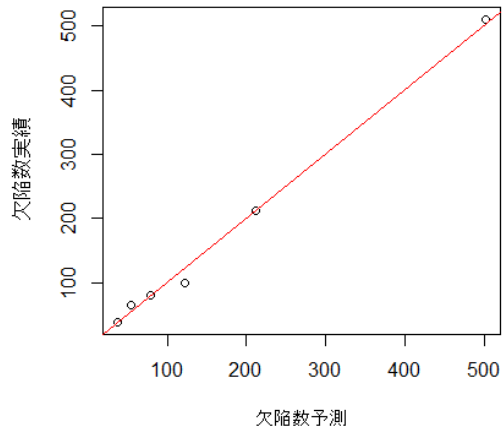


図 9：テスト欠陥数予測と実績

### 3.4 前倒し効果の可視化

影響評価モデルを用いて、2.3 節の実現したいことで述べたグラフを作成し、変更規模、前倒し率、レビュー欠陥数の関係を可視化する。

#### (1) 上流工程終了時におけるアクションの検討

レビュー欠陥数の実績値から前倒し率を予測するグラフを導出する。影響評価モデルの右辺は 2 つの変数を含んでいる。上流工程終了時点で変更規模はほぼ確定することになるため、これを代入する（ここでは仮に 50 KLOC とした）。すると右辺はレビュー欠陥数を変数とする指数関数となる。レビュー欠陥数を変動させた際のテスト欠陥数を求め、更に 3.2 節の定義から前倒し率を算出してプロットすると図 10 を得る。上流工程終了時点のレビュー欠陥数実績から、プロジェクト終了時の前倒し率が予測可能となり、目標値に対して差異が発生した場合には追加レビューを検討する、もしくはテスト戦略の修正を図るといったアクションの検討が可能である。

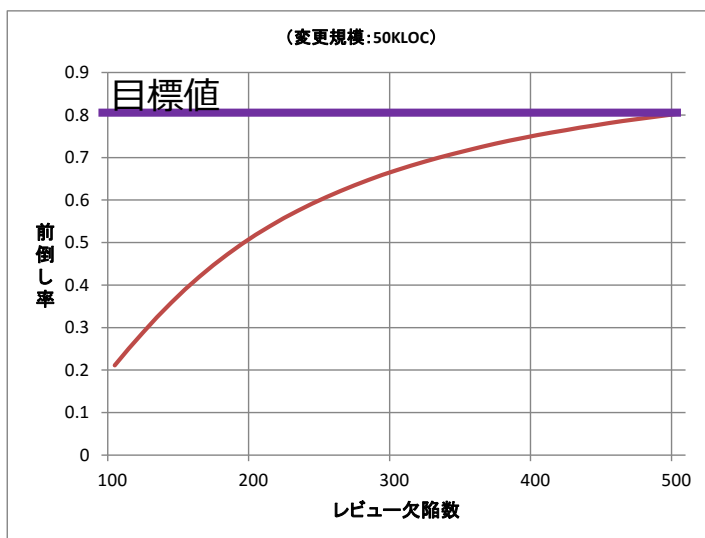


図 10：レビュー欠陥数からの予測



## (2) プロジェクト開始時におけるレビュー目標の設定

変更規模と前倒し率からレビューでの欠陥摘出数の目標値を設定するためのグラフを導出する。まず基準としたい前倒し率を設定する（ここでは60%および80%とした）。前倒し率の定義よりテスト欠陥数をレビュー欠陥数を用いて表現し、影響評価モデルの左辺をこれに置き換える。変更規模に任意の値を代入し、レビュー欠陥数を変動させて両辺の値が一致する点を求める。これを変更規模を変えて繰り返してプロットすると図11を得る。プロジェクト開始時に、変更規模の見積り値と前倒し率の目標値から、レビュー工程で摘出すべき欠陥数を算出することが可能である。

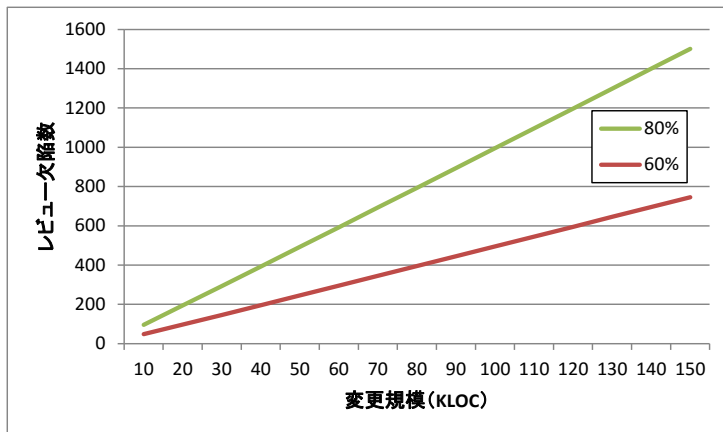


図11：変更規模および前倒し率からのレビュー欠陥数見積り

## 4. 考察

今回の対象データ数は6であり、統計解析による結果について信頼性が十分でないことは認めざるを得ない。まずは今回構築したモデルを足掛かりとして、今後のプロジェクトなどでデータを増やし、モデルを洗練させていく必要がある。一方、途中工程データの欠落と同様に、プロジェクト数が少ない中で改善活動を迫られることは少なくない。そのような場合でも影響評価モデルを構築し、改善活動に貢献できることを示した。

3.3節において、途中工程データが欠落した状況においても影響評価モデルの構築が可能であることを述べた。これを検証するには、全工程データが揃っているデータ群を用意し、全工程データを用いた場合と、途中工程データを除いた場合とで精度を算出し、その誤差を確認すればよい。

## 5. 結論

本研究では、途中工程のデータが欠落している状況における、上流でのレビュー活動が下流のテスト工程での品質に与える影響を定量的に示した影響評価モデルを構築した。その過程で、変更規模のみでは説明できなかったプロジェクトも、前倒し率を導入することによりモデルに反映されることを示した。更に、構築したモデルから、前倒し率、レビュー欠陥数、変更規模の関係を可視化し、上流工程終了時とプロジェクト開始時に活用できるよう工夫した。

モデル構築に前倒し率を用いることにより、下流工程を意識したレビュー活動を促すことを狙いとしている。また、前倒し率は変更規模等による影響を受けにくく、プロジェクトを跨っての観察に適している。前倒し率の改善に取り組むことで、プロジェクト単発の改善活動から脱却し、継続的な活動へ進展することが期待される。

## 謝辞

本論文執筆にあたり、お忙しい中にも惜しみなく相談に乗って下さった、小室 睦氏、小池 利和氏に心から感謝致します。

## 参考文献

- [1] 誉田 直美、ソフトウェア品質会計—NECの高品質ソフトウェア開発を支える品質保証技術、日科技連出版社、2010
- [2] 小室 睦、進化・発展的アプローチによるソフトウェアプロセス改善に対する予測モデル構築、Proceedings of FORCE2011, 2011
- [3] 小室 睦・薦田 憲久、ピアレビューデータに基づく品質予測モデル、電子情報通信学会論文誌 D. Vol. J94-D, No.2, 2011
- [4] 野中 誠・小池 利和・小室 睦、データ指向のソフトウェア品質マネジメント、日科技連出版、2012
- [5] Michael J. Crawley、統計学：Rを用いた入門書 第2版、共立出版、2016