

# 生産性要因に基づいて層別されたソフトウェア開発工数見積りモデル



角田 雅照 †



門田 暁人 †



松本 健一 †



高橋 昭彦 ††††

重回帰分析により工数見積りモデルを構築する場合、何らかの基準に基づいてプロジェクトを層別（分類）したデータセットを作成し、分類ごとに見積りモデルを構築することが望ましい。本論文では、生産性と関連が強い（寄与率の大きな）プロジェクト特性（開発言語、アーキテクチャ等）を用いて層別を行うことを提案する。財団法人 経済調査会が収集した153プロジェクトより抽出した、43プロジェクトの実績データを分析し、規模当たり要員数（平均要員数 / FP）の寄与率が大きい（寄与率38%）ことが分かった。一般に見積りモデル利用時には（規模当たり）要員数は確定していないが、規模当たり要員数を概算値（大、中、小の3段階）で決めることは比較的容易である。規模当たり要員数を用いてプロジェクトを3つに層別した結果、見積り精度が大きく改善する（要員数概算値の相対誤差平均が50%の場合、工数見積り値の相対誤差中央値が56.6%から35.9%になる）ことが分かった。

## Software Development Effort Estimation Models Stratified by Productivity Factors

Masateru Tsunoda †, Akito Monden †, Ken-ichi Matsumoto † and Akihiko Takahashi ††††

When building effort estimation models by multiple regression analysis, it is preferable to stratify software projects according to some criteria, and to build models for each stratified projects. In this paper, we propose to stratify projects by project attributes (programming language, and architecture et al.) which have strong relationships to productivity when building effort estimation models. We analyzed 43 projects data selected from 153 projects collected by Economic Research Association, and concluded that team size per project size had strongest relationship to productivity (variance explained was 38%). Although in general size per project size is not fixed when estimating effort, it is not very difficult that size per project is fixed approximately. We stratified projects by team size per project size and built models for each stratified projects. As a result, the accuracy of the models showed better performance than a model built from whole projects; and, the median of relative error of estimated effort was improved from 56.6% to 35.9% when the average of relative error of estimated team size was 50%.

### 1 はじめに

ソフトウェア開発プロジェクトの初期段階において開発工数を見積もることは、プロジェクト完遂に必要な資源の確保や、スケジュール管理を行う上で重要である。そのために、過去のソフトウェア開発プロジェクトの実

績データを見積りの根拠に用いる定量的見積り手法が数多く提案され、用いられてきた[BOEHM1981][SHEPPERD1997][SRINIVASAN1995][TSUNODA2005]。中でも、プロジェクトのメトリクス（開発規模等）を説明変数として用い、目的変数である開発工数との関係を一次式で表現する重回帰モデルを用いた見積りは、ソフトウェアツールが普及しており適用が容易なことから、広く実施されてきた[BRIAND2000][WALSTON1977]。

† 奈良先端科学技術大学院大学情報科学研究科, Graduate School of Information Science, Nara Institute of Science and Technology  
†† 財団法人経済調査会, Economic Research Association  
††† 産業技術大学院大学産業技術研究科, Graduate School of Industrial Technology, Advanced Institute of Industrial Technology

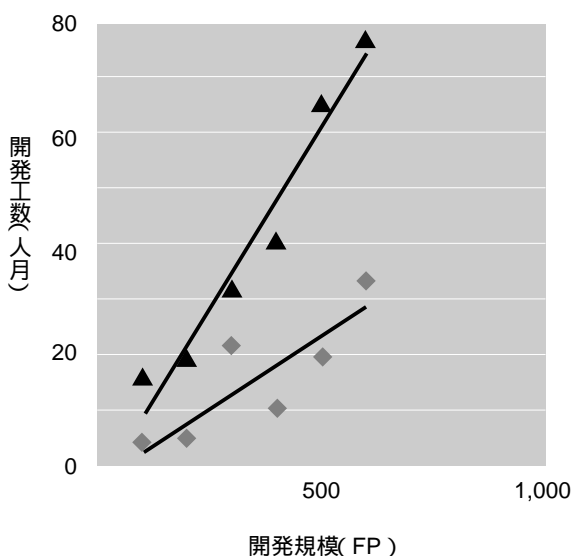


図1 プロジェクトの分類ごとに構築された見積りモデルの例

表1 生産性の統計量 (FP / 開発工数)

| 平均値  | 中央値  | 標準偏差 | 最小値 | 最大値  |
|------|------|------|-----|------|
| 15.5 | 12.8 | 10.9 | 1.3 | 48.7 |

一般に、重回帰分析により工数見積りモデルを構築する際には、多様なプロジェクトを含むデータセットをそのまま用いるのではなく、何らかの基準に基づいてプロジェクトを層別（分類）したデータセットを作成し、分類ごとに見積りモデルを構築することが望ましい [HOMMA2004]。図1はプロジェクトの分類ごとに構築された見積りモデルの（単純化した）例である。各点は1つのプロジェクトを表す。この例では、2つのタイプのプロジェクトが存在し、それぞれの開発規模（FP）と開発工数（人月）の関係を直線によりモデル化している。図に示すように、異なるタイプのプロジェクトが存在する場合、1つのモデル（直線）を当てはめて見積りモデルを構築したとしても、見積り誤差の小さなモデルは得られない。このため、プロジェクトを分類するための系統的な方法が求められる。

ここで図1を詳細に観察すると、2つの直線は傾きが大きく異なっており、この傾きの違いは生産性（開発規模（FP）/開発工数（人月））の違いであり、生産性が大きく異なるプロジェクトが混在していることが分かる。見積り誤差の小さなモデルを得るためには、生産性の違いに基づいてプロジェクトを分類（層別）し、分類ごとに

見積りモデルを作る必要がある。ただし、生産性は開発規模/開発工数で表され、開発工数はプロジェクトが完了するまで確定しないため、工数を見積る時点では生産性は未知である。

本論文では、見積りを行う時点において決定可能なプロジェクト特性（開発言語、アーキテクチャ等）のうち、生産性との関連が強い（寄与率の大きな）特性を用いて層別を行うことを提案する。これにより、見積りを行う時点において決定しているプロジェクト特性に基づいて、利用する見積りモデルを決定することが出来る。

本論文では、財団法人 経済調査会が収集した43プロジェクトの実績データを用いて、各プロジェクト特性の（生産性に対する）寄与率を求める。そして、寄与率の大きなプロジェクト特性を用いてプロジェクトを層別し、見積りモデルを構築する。層別の効果を評価するために、層別しない場合の見積りモデルと見積り精度を比較する。

以降、第2節では分析対象のデータについて説明し、第3節では生産性とプロジェクト特性との関連について分析する。第4節では生産性と関連の強いプロジェクト特性に基づいてプロジェクトを層別し、モデルを構築した場合の見積り精度を比較する。第5節では実験結果に対する考察を行い、第6節では関連研究について説明し、第7節でまとめを述べる。

## 2 分析対象データ

分析に用いたデータは、日本ファンクションポイントユーザ会FP法利用検討会（JFPUG）及び奈良先端科学技術大学院大学との共同調査として、財団法人 経済調査会が平成18年度に110社から収集した、153プロジェクトの実績値である。データには、従業員数4人の企業から従業員数約40,000人の企業のプロジェクトまで含まれており、非常に多様であるという特徴がある。プロジェクトごとに、約140個の変数（プロジェクト特性）が記録されている。ただし、変数によっては欠損値（値が未記録であること）が多く含まれている場合があった。

分析対象のプロジェクトの条件を揃えるため、下記の条件に従ってデータを抽出した。

- ・新規開発のプロジェクト（改造のプロジェクトは分析

表2 分析対象のプロジェクト特性

| プロジェクト特性    | 種別   | 件数 | 詳細説明  |
|-------------|------|----|---|
| 発注者分類       | カテゴリ | 42 | 府省庁 地方自治体 民間企業(親会社) 民間企業(親会社以外)   |
| システム構成      | カテゴリ | 42 | クライアントサーバシステム Web系システム メインフレームシステム  |
| 適用分野        | カテゴリ | 40 | 事務系 制御系   |
| 適用業種        | カテゴリ | 40 | 建設業 製造業 電気・ガス・熱供給・水道業 情報通信業 運輸業 卸売・小売業, 金融・保険業 不動産業 飲食店・宿泊業 医療・福祉 教育・学習支援業 公務 |
| サーバOS       | カテゴリ | 37 | UNIX Linux Windows NT 系   |
| データベース      | カテゴリ | 42 | Microsoft SQL Server Oracle その他(オープンソース) その他(非オープンソース)                        |
| 開発技法        | カテゴリ | 39 | 構造化 DOA オブジェクト指向  |
| 開発作業標準      | カテゴリ | 39 | 使用した 使用しなかった  |
| 主開発言語       | カテゴリ | 32 | 使用割合が50%以上の開発言語<br>ASP ASP.NET COBOL Delphi Java SQL VB VB.NET                |
| 平均要員数       | 数値   | 40 | 平均要員数 = $\frac{\text{開発工数}}{\text{開発期間}}$                                     |
| 開発期間        | 数値   | 40 | 開発期間の実績値  |
| FP          | 数値   | 43 | 未調整ファンクションポイントの実績値  |
| 規模当たり期間     | 数値   | 40 | 規模当たり期間 = $\frac{\text{平均期間}}{\text{FP}}$                                     |
| 規模当たり要員数    | 数値   | 38 | 規模当たり要員数 = $\frac{\text{平均要員数}}{\text{FP}}$                                   |
| 基本設計比率      | 数値   | 28 | 全工程に対する基本設計工程の割合  |
| 詳細設計比率      | 数値   | 35 | 全工程に対する詳細設計工程の割合  |
| PG設計製造比率    | 数値   | 37 | 全工程に対するプログラム設計・製造工程の割合  |
| ソフトウェアテスト比率 | 数値   | 35 | 全工程に対するソフトウェアテスト工程の割合   |
| システムテスト比率   | 数値   | 31 | 全工程に対するシステムテスト工程の割合   |

から除外した)。

- ・基本設計, 詳細設計, プログラム設計・製造, ソフトウェアテスト, システムテストの工程比率の合計が100%のプロジェクト。もしくは, 各工程の比率は欠損しているが, 開発工程の範囲が基本設計からシステムテストまでのプロジェクト(一部の工程しか実施していないプロジェクトは分析から除外した)。
- ・FP(ファンクションポイント), 開発工数が計測されているプロジェクト(生産性を計算するためにはFP, 開発工数が計測されている必要がある)。

データを抽出した結果, 分析対象のプロジェクトは43件となった。分析対象のプロジェクトのうち37件は事務系, 4件は制御系, 残りはその他のシステム開発であり, 組込み系システム開発のプロジェクトは含まれていなかった。1件のプロジェクトを除き, 開発方法論はウォーターフォールであった。

本論文では, FPを基本設計からシステムテストにおける開発(総)工数(人月)で割った値を生産性と定義した。表1に抽出した43件のプロジェクトにおける生産性

の統計量を示す。生産性の第1四分位数(生産性を高い順に並べた場合の下位25%)と第3四分位数(生産性を高い順に並べた場合の上位25%)を比較すると, 約3.3倍の差があった。

分析を行ったプロジェクト特性を表2に示す。欠損値の多いプロジェクト特性や, ほとんどのプロジェクトで同一の値となっていたプロジェクト特性は分析から除外した。ただし, 分析対象の変数にも欠損値が含まれているため, 表2に示すように, 変数の値が記録されているプロジェクト数はそれぞれ異なる。

データに含まれていない平均要員数, 規模当たり要員数, 規模当たり期間を新たに定義した。平均要員数は開発工数を開発期間で除したものである。一般に, 開発要員数と開発期間は開発規模(FP)に比例して大きくなるため, 規模でそれぞれを除すことにより正規化した, 規模当たり要員数と規模当たり期間を定義した。

なお, 一般に開発要員数は見積り工数から決定することが多く, 工数見積りモデルを利用する際には(開発工数が未確定の時点では)開発要員数や規模当たり要員数が確定していないことが多い。ただし, モデルを利用す

際には、開発要員数や規模当たり要員数の正確な値を決定する必要はなく、後述するように概算値（大，中，小の3段階）を決定するだけでモデルを利用することが可能であるため、分析対象に含めた。

### 3 生産性の分析

本論文では、以下の手順で分析を行う。

**手順1** 分散分析の寄与率を用いて、生産性と関連の強いプロジェクト特性を明らかにする。

**手順2** 生産性と関連の強いプロジェクト特性に基づいてプロジェクトを層別し、工数見積りモデルを作成する。

本節では、生産性と関連の強いプロジェクト特性を明らかにし、第4節においてプロジェクトを層別した工数見積りモデルを作成する。

#### 3.1 生産性と関連の強い変数の特定

生産性と関連の強いプロジェクト特性を特定するために、一元配置分散分析の有意確率と（調整済み）寄与率を求め、生産性と各プロジェクト特性との関連の有無と強さを分析した。一元配置分散分析は、カテゴリデータと数値データの変数の関連の有無と強さを分析する手法であり、有意確率が有意水準未満になった場合、そのプロジェクト特性と生産性は統計的に見て関連があることを示す（本論文では有意水準を10%とした）。また、寄与率は関連の強さを表す。寄与率は0%（場合によっては負の値となる場合もある）から100%の値をとり、寄与率が大きいほどプロジェクト特性と生産性の関連が強いことを示す。寄与率は以下の式によって求める[WINER1991]。

$$\text{寄与率} = \frac{SSB - (k - 1)MSE}{SST + MSE} \quad (1)$$

ここで、 $SSB$ は群間平方和、 $SST$ は総平方和、 $MSE$ は平均二乗誤差、 $k$ はグループ数を表す。寄与率はデータ全体のばらつき（ $SST$ ）に対し、各要因のばらつき（ $SSB$ ）がどの程度寄与しているかを示している。 $MSE$ はデータの分布の偏りを補正するために用いられており、 $k$ は各要因のカテゴリ数の違いを補正するために用いられている。

一元配置分散分析に当たり、すべてのプロジェクト特性をカテゴリデータの変数として扱うため、数値データのプロジェクト特性をカテゴリデータに変換した。これ

により、平均要員数等の数値データのプロジェクト特性と、開発言語等のカテゴリデータのプロジェクト特性の、生産性に対する関連の強さが比較可能となる。各特性の第1四分位数以下（値を大きい順に並べた場合の下位25%以下）の値を「低位」カテゴリとした。同様に、各特性の第3四分位数以上（値を大きい順に並べた場合の上位25%以上）の値を「高位」カテゴリとし、残りの値を「中位」カテゴリとした。これらのカテゴリの閾値を表3に示す。なお、カテゴリデータに変換することにより、本来あったはずの関連性が消える危険性があるため、数値データのままで順位相関係数を算出し、離散化を行った場合と同様の傾向が見られるかどうかを確かめる。

表3 各プロジェクト特性のカテゴリと閾値との対応

| プロジェクト特性    | 「低位」カテゴリ | 「高位」カテゴリ |
|-------------|----------|----------|
| 平均要員数       | 2.6      | 14.1     |
| 開発期間        | 5.7      | 10.3     |
| FP          | 262.5    | 1,494.0  |
| 規模当たり期間     | 0.007    | 0.018    |
| 規模当たり要員数    | 0.005    | 0.012    |
| 基本設計比率      | 12.9     | 20.3     |
| 詳細設計比率      | 14.9     | 25.0     |
| PG設計製造比率    | 30.0     | 40.0     |
| ソフトウェアテスト比率 | 15.0     | 25.0     |
| システムテスト比率   | 10.0     | 20.0     |

表4 生産性に対する関連の有無及び関連の強さ

| プロジェクト特性    | 寄与率 | 有意確率 |
|-------------|-----|------|
| 規模当たり要員数    | 38% | 0%   |
| 平均要員数       | 31% | 1%   |
| 規模当たり期間     | 19% | 1%   |
| 開発期間        | 9%  | 6%   |
| 適用業種        | 7%  | 27%  |
| FP          | 7%  | 9%   |
| サーバOS       | 6%  | 13%  |
| 発注者分類       | 5%  | 19%  |
| データベース      | 3%  | 27%  |
| 適用分野        | 2%  | 21%  |
| 開発作業標準      | -1% | 37%  |
| 開発技法        | -1% | 43%  |
| 基本設計比率      | -3% | 56%  |
| 主開発言語       | -4% | 57%  |
| システム構成      | -4% | 80%  |
| システムテスト比率   | -5% | 71%  |
| PG設計製造比率    | -5% | 88%  |
| 詳細設計比率      | -6% | 92%  |
| ソフトウェアテスト比率 | -6% | 97%  |



一元配置分散分析の結果を表4に示す。関連があったプロジェクト特性の有意確率を斜体としている。5つのプロジェクト特性（規模当たり要員数，平均要員数，規模当たり期間，開発期間，FP）が生産性と関連があり（有意水準が10%未満となり），残りのプロジェクト特性（適用業種，サーバOS，発注者分類等）は生産性と関連があるとは言えなかった。

生産性と関連があったプロジェクト特性のうち，規模当たり要員数の関連が最も強く（寄与率 38%），FP（開発規模）は弱い関連しか見られなかった（寄与率 7%）。直感的には，規模の大きなシステムは開発が難しい（すなわち生産性が低い）と考えられるが，分析においては，開発規模（FP）は生産性との関連があまり強くなかった。さらに，規模当たり要員数の寄与率が平均要員数の寄与率よりも大きかったことから，プロジェクトの生産性に対しては，単純な開発要員数の多寡よりも，開発規模に比して適切な開発要員数かどうか重要であると言える。

数値データのプロジェクト特性と生産性との順位相関係数を表5に示す。規模当たり要員数，平均要員数，開発期間，FPの順に生産性との関連が強くなり，相関も有意であり，各工程比率は生産性との関連が弱く，相関も有意ではなかった。これは，表4の一元配置分散分析の結果と同様の傾向である。よって，カテゴリデータに変換することにより，本来あったはずの関連性が消えていることはないと考えられる。なお，規模当たり期間と生産性との相関係数が小さくなっていったが，これは規模当たり期間と生産性の関係が，3.2.3節で後述するように単調増加または単調減少でないためである。

### 3.2 生産性と関連が強い要因の個別分析

3.1節で生産性と統計的に関連が見られた5つのプロジ

表5 数値データのプロジェクト特性と生産性との順位相関係数

| プロジェクト特性    | 相関係数  | 有意確率 |
|-------------|-------|------|
| 規模当たり要員数    | -0.65 | 0%   |
| 平均要員数       | -0.59 | 0%   |
| 開発期間        | -0.36 | 2%   |
| FP          | -0.28 | 7%   |
| 基本設計比率      | -0.19 | 34%  |
| 詳細設計比率      | 0.09  | 62%  |
| 規模当たり期間     | 0.08  | 62%  |
| ソフトウェアテスト比率 | 0.06  | 71%  |
| システムソフト比率   | -0.01 | 96%  |
| PG設計製造比率    | 0.00  | 98%  |

ジェクト特性（規模当たり要員数，平均要員数，規模当たり期間，開発期間，FP）について，箱ひげ図を作成し，プロジェクト特性の値と生産性との関係を分析した。以降に詳細を述べる。

#### 3.2.1 規模当たり要員数

規模当たり要員数のカテゴリ別（規模当たり要員数=「低位」，「中位」，「高位」）の生産性の箱ひげ図を図2に示す。それぞれの箱の太線は生産性の中央値を表す。丸印は生産性の外れ値を表す。外れ値は箱の上端（もしくは下端）から，四分位範囲（箱の長さ）の1.5倍以上離れた値である。星印は極値（極値は箱の上端（もしくは下端）から，四分位範囲（箱の長さ）の3倍以上離れた値）を表す。

分析の結果，規模当たり要員数が大きいプロジェクト程，生産性が低くなっていった。低位カテゴリの生産性の中央値は，高位カテゴリの生産性の中央値の約3.7倍であった。各カテゴリの生産性の違いを統計的に確かめると（マン・ホイットニーのU検定を用い，有意水準は10%と

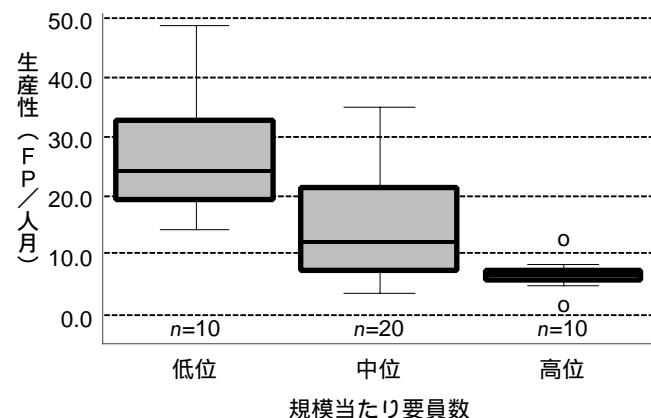


図2 規模当たり要員数と生産性の関係

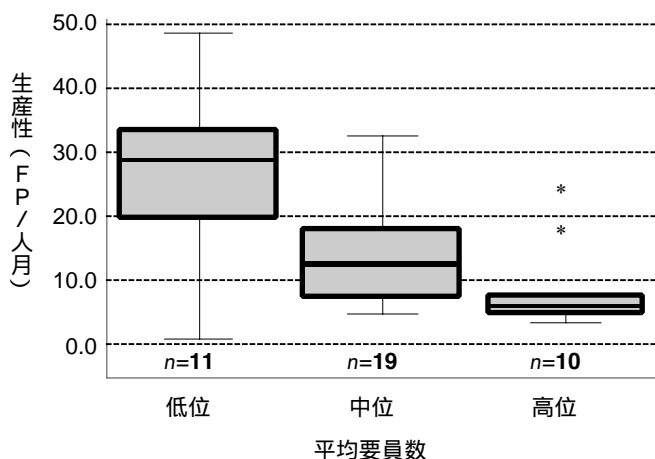


図3 平均要員数と生産性の関係

した),すべてのカテゴリ間で差があると言えた。開発期間が短く制限されている場合等,開発規模に比して開発要員を多く投入する場合(本論文で用いたデータでは1,000FP当たりの平均要員数が12人以上),とくに生産性が低くなると考えられる。規模当たり要員数と生産性との関連が大きかった理由は,規模当たり要員数が増えるとき,開発要員間のコミュニケーションのオーバーヘッドが大きくなるためであると推測される。

### 3.2.2 平均要員数

図3に平均要員数によって分類したプロジェクトの生産性の箱ひげ図を示す。図より,平均要員数が増加すると,生産性が低下していることが分かる。低位カテゴリのプロジェクトの生産性の中央値は,高位カテゴリのプロジェクトの約4.5倍であった。各カテゴリの生産性の違いを統計的に確かめると,すべてのカテゴリ間で差があると言えた。平均要員数が多い(14人以上)プロジェクトでは,生産性は低くなると考えられる。

### 3.2.3 規模当たり期間

図4に規模当たり期間によって分類したプロジェクトの生産性の箱ひげ図を示す。規模当たり期間が中位のプロジェクトの生産性が高くなっていた。中位カテゴリの生産性の中央値は,低位カテゴリの生産性の中央値の約2.7倍であった。高位カテゴリの生産性の中央値は,低位カテゴリの生産性の中央値の約1.1倍であり,ほとんど差が無かった。各カテゴリの生産性の違いを統計的に確かめると,中位カテゴリとその他のカテゴリ間で差があると言えた。開発規模に比して開発期間を短くすると(1,000FP当たり開発期間が7ヵ月以下)生産性は低下するが,逆に長くしすぎても(1,000FP当たり開発期間が

18ヵ月以上)生産性は低くなると考えられる。

### 3.2.4 開発期間

図5に工期に応じて3つに分類した場合の,プロジェクトの生産性の箱ひげ図を示す。開発期間が低位の(開発期間が短い)プロジェクトの生産性は,開発期間が中位や高位のプロジェクトの生産性よりも比較的高くなっていった。低位カテゴリの生産性の中央値は,高位カテゴリの生産性の中央値よりも約3.1倍大きかった。各カテゴリの生産性の違いを統計的に確かめると,低位カテゴリと高位カテゴリの間で差があると言えた。

ただし,各カテゴリの生産性の中央値を比較すると差があるが,四分位範囲(箱の位置)の差は比較的小さかった。すなわち,開発期間が低位の(短い)プロジェクトが,開発期間が高位の(長い)プロジェクトよりも生産性が高いということが比較的発生しやすいということであり,開発期間がプロジェクトの生産性に与える影響が小さいことが分かる。0の結果と合わせると,プロジェクトの生産性に対しては,単純な開発期間の長短より

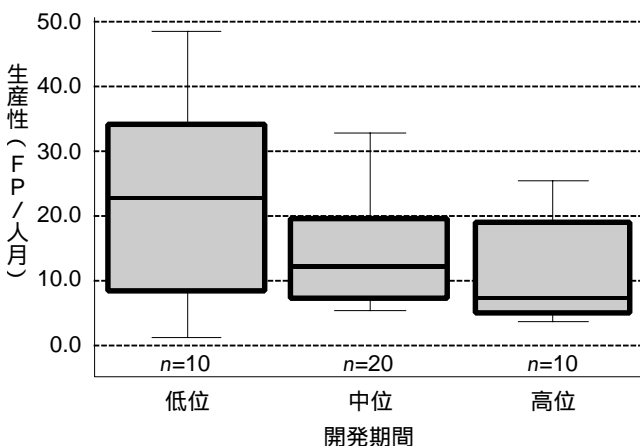


図5 開発期間と生産性の関係

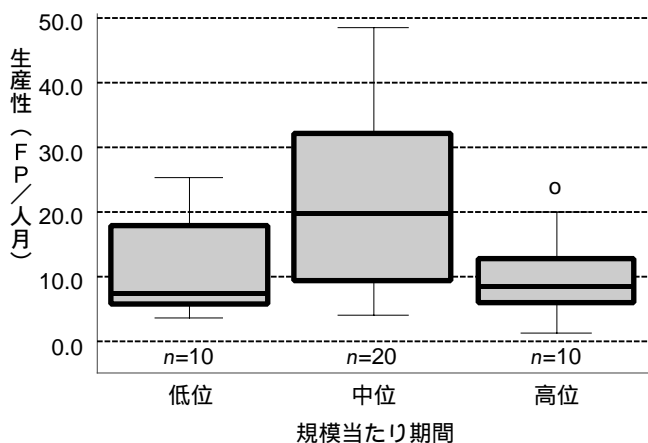


図4 規模当たり期間と生産性の関係

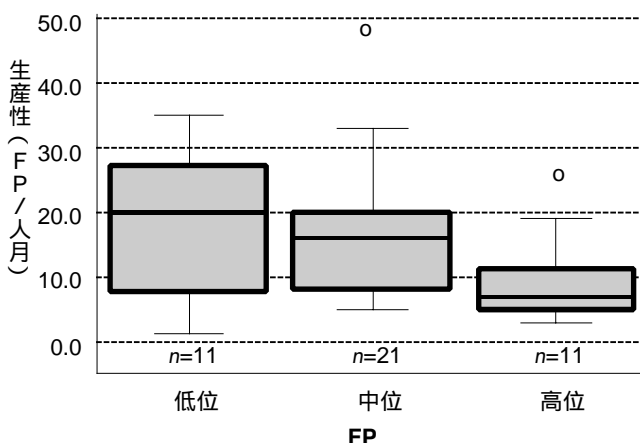


図6 FPと生産性の関係

も、開発規模に比した開発期間の長短のほうに影響していると考えられる。

### 3.2.5 FP

FP（開発規模）によって分類したプロジェクトの生産性の箱ひげ図を図6に示す。低位カテゴリのプロジェクトの生産性の中央値は、高位カテゴリのプロジェクトの生産性の中央値よりも約3.0倍大きかった。各カテゴリの生産性の違いを統計的に確かめると、高位カテゴリと他のカテゴリとの間で差があると言えた。図からも、FPが高位の（開発規模が大きい）プロジェクトの生産性は比較的低くなっており、FPが中位や低位のプロジェクトでは生産性の差は小さいことが分かる。システムの規模が特に大きい（本論文で用いたデータではFPが1,500以上）プロジェクトの場合は、生産性が低下しやすい傾向があると言える。

## 4 工数見積りモデルの構築

第3節の分析結果より、規模当たり要員数が最も生産性と関連が強いことが分かった。そこで、規模当たり要員数に基づいてプロジェクトを層別して工数見積りモデルを構築する。そして、従来から生産性と関連が強いことが明らかとなっている平均要員数に基づいてプロジェクトを層別してモデルを構築した場合との工数見積りの精度を比較する。また、開発規模により層別してモデルを構築することが一般に行われているため、この場合との見積り精度の比較を行うと共に、プロジェクトを層別せずにモデルを構築した場合との見積りの精度の比較も行う。

実験では、モデル構築時に用いるプロジェクト（学習データ）を規模当たり要員数が小（規模当たり要員数 0.005）、中（ $0.005 < \text{規模当たり要員数} < 0.012$ ）、大（規模当たり要員数 0.012）に層別し、各層ごとに工数見積りモデルを構築した。そして、例えば見積り対象プロジェクトの規模当たり要員数が大（規模当たり要員数が0.012以上）となるとき、規模当たり要員数が大のプロジェクトを用いて作成したモデルを用いて見積りを行った。平均要員数に基づいてプロジェクトを層別してモデルを構築し、見積り場合も同様の手順で行った。プロジェクトを層別する際の閾値は表3の値を用いた。

なお、開発要員数は見積り工数から決定することが多

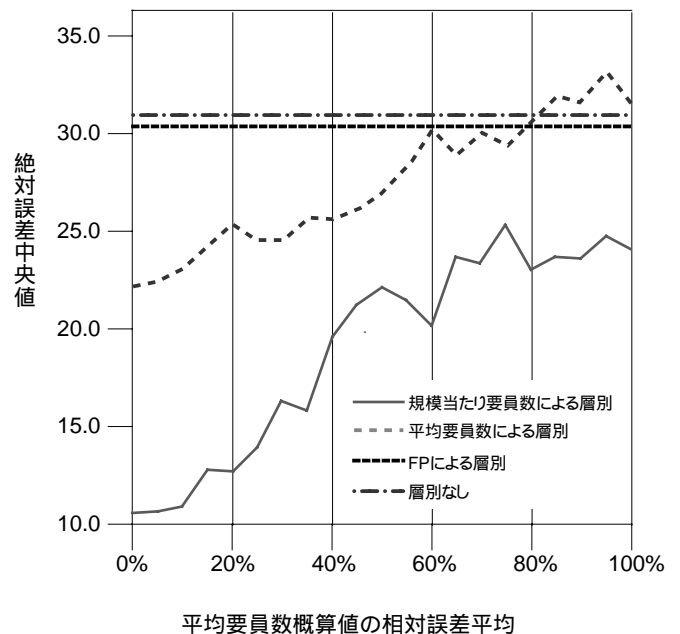


図7 見積り工数の絶対誤差中央値と平均要員数の概算値の相対誤差平均との関係

いため、工数見積りモデルを利用する際には（開発工数が未確定の時点では）開発要員数が確定していないことが多い。この場合、開発要員数を概算で決定し、開発要員数の概算値とFPにより規模当たり要員数を求め、規模当たり要員数に応じて使用するモデルを決定する必要がある。ただし、開発要員数の概算値には誤差があると考えられるため、使用するモデルを誤って選択する可能性がある。

そこで本論文では、開発要員数の概算値に平均 $n\%$ の相対誤差が含まれると仮定し、モデルの選択を誤った場合に、工数見積り精度がどの程度変化するかを実験により確かめた。開発要員数の相対誤差は正規分布に従っていると仮定し、開発要員数の相対誤差平均 $n$ を5%刻みで変化させると共に、プロジェクトごとに誤差を含んだ平均要員数の概算値を20回生成し、工数見積りモデルの精度を確かめた。

工数見積りモデルは回帰分析により作成した。(重)回帰分析は、工数見積りモデルの構築に広く用いられている手法である。モデルの説明変数はFPのみとした。表2のプロジェクト特性のうち、工数見積り時に確定しており、かつ工数見積りに役立つと考えられる適用業種と主開発言語を説明変数に含めて工数見積りモデルを構築す

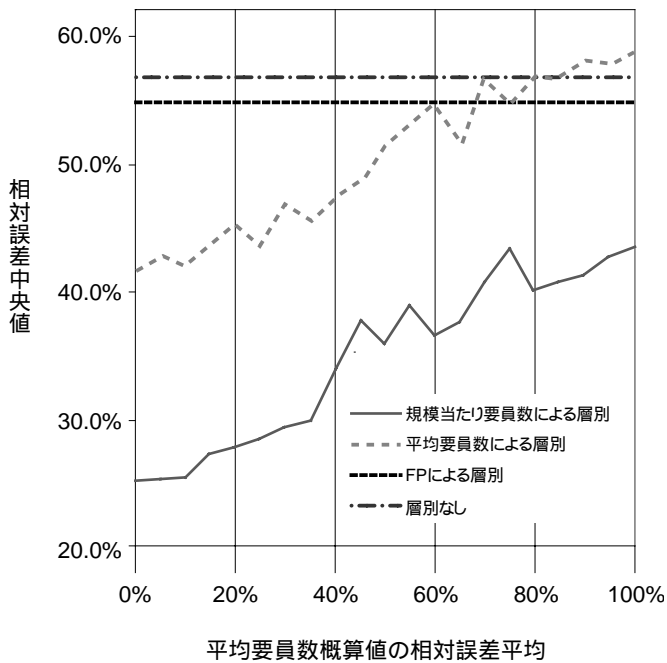


図8 見積り工数の相対誤差中央値と平均要員数の概算値の相対誤差平均との関係

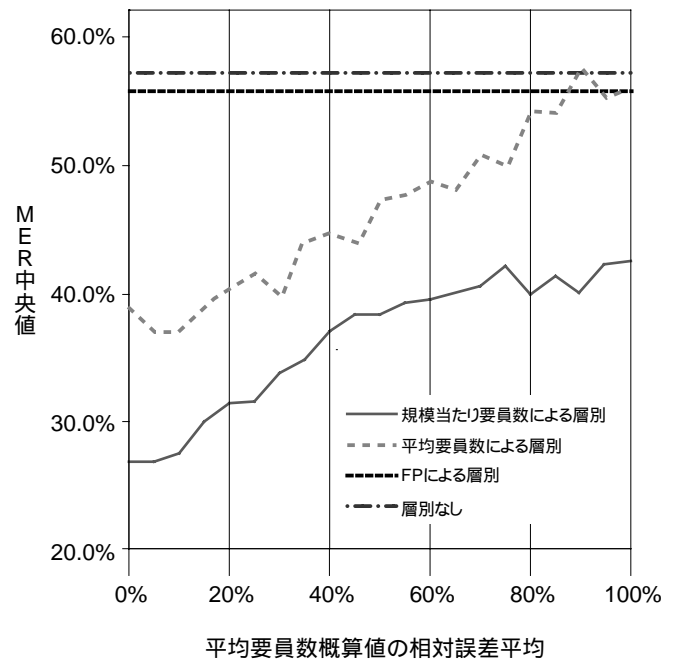


図9 見積り工数のMER中央値と平均要員数の概算値の相対誤差平均との関係

る予備実験を行ったが、工数見積りに役立つとは言えなかった（ステップワイズ重回帰分析を行った結果，説明変数として採用されなかった）ため，これらを説明変数に含めなかった．また，開発期間は工数見積りモデルの利用時（開発工数が未確定のとき）に確定していない場合があるため，開発期間と規模当たり開発期間を説明変数に含めなかった．

正規Q-Qプロット[TANAKA1995]により変数の分布を確かめると，FPと開発工数が正規分布していなかったため，それぞれを対数変換してから回帰分析を行った．また，全プロジェクトに対して回帰分析を行い，CookのDが2.73となったプロジェクト1件を外れプロジェクトとみなし，これを除外して実験を行った．CookのDは外れケース（プロジェクト）を特定する手法であり[TANAKA1995]，一般にCookのDが1を超えているケースは外れケースとみなされる．

実験ではリーブワンアウト法を用いた．リーブワンアウト法とは，全プロジェクトのうち1つのプロジェクトをモデル評価用のテストデータ，残りのプロジェクトをモデル構築用の学習データとすることを，すべてのプロジェクトに対して行うことにより，モデルを評価する方法である[HASTIE2001]．

モデルの精度の評価基準として，一般的に用いられる絶対誤差（ $|\text{見積り値} - \text{実測値}|$ ），相対誤差（絶対誤差 / 実測値）[CONTE1986]，MER（絶対誤差 / 予測値）[KITCHENHAM2001]を用いた．これらの値が小さいほど，モデルの精度が高いことを示す．

図7から図9に各見積りモデルの精度を示す．層別なしの場合，FPにより層別した場合は各評価指標の中央値である．規模当たり要員数により層別した場合，平均要員数により層別した場合は，各評価指標の中央値20個の平均値（プロジェクトごとに誤差を含んだ平均要員数の概算値を20回生成しているため）である．

規模当たり要員数によってプロジェクトを層別してモデルを構築した場合，平均要員数の概算値の相対誤差平均が100%の場合でも，平均要員数により層別して構築したモデルや，FPにより層別して構築したモデル，層別せずに構築したモデルよりも，絶対誤差，相対誤差，MERの各中央値が小さくなっていた．

平均要員数によってプロジェクトを層別してモデルを構築した場合，平均要員数の概算値の相対誤差平均が60%を超えると，絶対誤差，相対誤差の中央値が，FPにより層別した場合とほとんど変わらないか，大きくなっていた．更に，平均要員数の概算値の相対誤差平均が



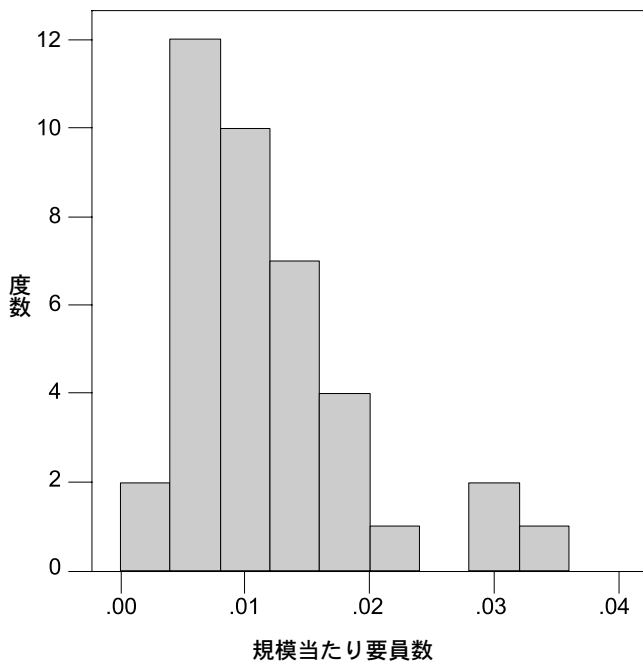


図10 規模当たり要員数の分布

85%を超えると、絶対誤差，相対誤差の中央値が，層別しない場合とほとんど変わらないか，大きくなっていった。

FPに基づいてプロジェクトを層別してモデルを構築した場合と，層別せずにモデルを構築した場合を比較すると，FPにより層別した場合の見積り精度がわずかに高いが，見積り精度に大きな違いは見られなかった。

これらの結果から，生産性に関連の強い規模当たり要員数によってプロジェクトを層別し，モデルを構築した場合，平均要員数の概算値の相対誤差平均が100%の場合でも，平均要員数により層別して構築したモデルや，FPにより層別して構築したモデル，層別せずに構築したモデルよりも見積り精度が高くなると言える。

## 5 考察

実験結果より，規模当たり要員数に基づいてプロジェクトを層別し，工数見積りモデルを作成することにより，見積り精度が高まることが分かった。

一般に，規模当たり要員数の大小は，工期の厳しさ[MONDEN2007]に影響されると考えられる。非常に短い開発期間で開発を行う場合，納期を守るために多くの要

員を投入せねばならず，規模当たり要員数は大きくなる。一方，開発期間に余裕がある場合にはより少人数で開発出来るため規模当たり要員数は小さくなる。このことから，本実験の結果は，工期の厳しさに応じてモデルを作ることの重要性を示唆している。逆に言えば，工期が厳しいプロジェクトとそうでないプロジェクトを混在させてモデルを作ることの危険性を示しているとも言える。

なお，開発プロジェクトの計画立案に際しては，規模当たり要員数には上限値がある（すなわち，工期を厳しくするにも限度がある）ことに留意されたい。図10に規模当たり要員数の分布を示す。規模当たり要員数が0.02以上のプロジェクトは少なく，また規模当たり要員数が0.035以上のプロジェクトは1件しか存在しなかった。よって，工期が厳しい場合でも，規模当たり要員数を0.035以上にするのは難しいと考えられる。

次に，プロジェクト特性を構成するメトリクス間の関係について考察する。層別に用いたプロジェクト特性と生産性を開発工数=MM，開発期間=DDを用いて表すと，以下ようになる。

$$\text{生産性} = \frac{FP}{MM}$$

$$\text{平均要員数} = \frac{MM}{DD} = \frac{FP}{\text{生産性}}$$

$$\text{規模当たり要員数} = \frac{\text{平均要員数}}{FP} = \frac{MM}{DD \cdot FP} = \frac{1}{\text{生産性}}$$

すなわち，平均要員数，規模当たり要員数とも，定義の分母に生産性を含んでいることになるが，これらのメトリクスが生産性と負の相関を持つことは必ずしも自明ではない。例えば，生産性（FP / MM）の分子はFPであり，分母は開発工数（MM）となっていることから，生産性は分子のFPと正の相関があり，分母の工数とは負の相関があることは一見自明であるように思われる。しかし，実際に順位相関係数を調べると，FPと生産性との相関係数は - 0.28，工数と生産性との相関係数は - 0.60となっており，いずれも生産性と負の相関となっていた。すなわち，分子にあるから正の相関，分母にあるからといって負の相関になるとは一概に言えないことになる。

また，分析対象のプロジェクト特性の多くは，生産性を構成するメトリクスを含んでおり，生産性と何らかの

関連があることは一見自明であるが、生産性に対してプラスに働くのかマイナスに働くのかは、定義からは明らかではない。例えば、FP（開発規模）は生産性に対してマイナスに働くのは一見自明のようであるが、ソフトウェア開発以外の業種（製造業等）では、規模が大きいと量産効果によって生産性が向上し、プラスに働く場合もあり得る。同様の例としてバグ密度が挙げられる。バグ密度はバグ数／開発規模により定義されるが、開発規模がバグ密度に対してプラスに働くのかマイナスに働くのかは、定義からは明らかではない。

## 6 関連研究

これまでソフトウェア開発の生産性について、様々な国のデータリポジトリを用いた研究が行われており、開発要員数と生産性との関連が強いことを示した研究が幾つか存在する[BLACKBURN1996][JEFFERY1996][MAXWELL1996]。ただし、規模当たりの開発要員数と生産性との関連はこれまで明らかにされていなかった。本論文では、開発要員数よりも規模当たりの開発要員数のほうが、より生産性に強い影響を与えることを示した。

開発要員数を考慮したモデルとして、開発要員数をそのまま説明変数に用いたモデルが幾つか提案されている[JEFFERY2000][MAXWELL1999][MENDES2004]。ただし、工数見積りモデルの利用時（開発工数が未確定のとき）に開発要員数が確定していることは少ないため、モデルの利用可能機会は少ないと考えられる。また、規模当たり要員数を説明変数に用いたモデルや、規模当たり要員数により層別した工数見積りモデルは提案されていない。

## 7 まとめ

本論文では、生産性との関連が強いプロジェクト特性に基づいてプロジェクトを層別し、工数見積りモデルを構築することを提案した。分析により、規模当たり要員数は生産性との関連が強いことを明らかにした。そして、規模当たり要員数に基づいてプロジェクトを層別して構築したモデルは、平均要員数に基づいて層別して構築したモデルや、FPに基づいて層別して構築したモデル、層別を行わずに構築したモデルよりも、工数見積りの精度

が高いことを確かめた。今後の予定は、様々なデータセットを用いて提案方法の見積り精度を確認し、結果の信頼性をより高めることである。

謝辞

本研究の一部は、文部科学省「次世代IT基盤構築のための研究開発」の委託に基づいて行われた。

参考文献

- [BLACKBURN1996] J. Blackburn, G. Scudder, and L. Wassenhove : Improving speed and productivity of software development: a global survey of software developers, IEEE Trans. on Software Eng., Vol.22, No.12, pp.875-885, 1996
- [BOEHM1981] B.W. Boehm : Software Engineering Economics, Prentice Hall, New Jersey, 1981
- [BRIAND2000] L. Briand, T. Langley, and I. Wiecezorek : A replicated assessment and comparison of common software cost modeling techniques, Proc. 22nd IEEE International Conference on Software Eng., Limerick, pp.377-386, 2000
- [CONTE1986] S. Conte, H. Dunsmore, and V. Shen : Software Engineering Metrics and Models, Benjamin-Cummings, MenloPark, California, 1986
- [HASTIE2001] T. Hastie, R. Tibshirani, and J. Friedman : The Elements of Statistical Learning: Data Mining, Inference and Prediction, Springer-Verlag, New York, 2001
- [HOMMA2004] 本間周二 : CSKに見るFP法の実践 プロジェクト定量化で品質と見積り精度向上, 日経ITプロフェッショナル, pp.53-61, 2004-7
- [JEFFERY1996] R. Jeffery : Time-sensitive cost models in the commercial MIS environment IEEE Trans. on Software Eng., Vol.13, No.7, pp.852-859, 1996
- [JEFFERY2000] R. Jeffery, M. Ruhe, and I. Wiecezorek : A comparative study of two software development cost modeling techniques using multi-organizational and company-specific data Information and Software Technology, Vol.42, No.14, pp.1009-1016, 2000
- [KITCHENHAM2001] B. Kitchenham, L. Pickard, and S. MacDonell : What accuracy statistics really measure, IEE Proceedings - Software, Vol.148, Issue 3, pp.81-85, 2001
- [MAXWELL1996] K. Maxwell, L. Wassenhove, and S. Dutta : Software development productivity of European space, military, and industrial applications, IEEE Trans. on Software Eng., Vol.22, No.10, pp.706-718, 1996
- [MAXWELL1999] K. Maxwell, L. Wassenhove, S. Dutta : Performance evaluation of general and company specific models in software development effort estimation Management Science Vol.45, No.6, pp.787-803, 1999
- [MENDES2004] E. Mendes, and B. Kitchenham : Further comparison of cross-company and within-company effort estimation models for web applications, Proc. 10th International Software Metrics Symposium, pp.348-357, 2004
- [MONDEN2007] 門田暁人, 馬嶋宏, 増田浩, 羽田野尚登, 磯野聖, 内海昭, 菊地奈穂美, 服部昇, 細谷和伸, 森和美 : 工期の厳しさに関連する要因の分析, SEC journal, No.10, pp.48-53, 2007
- [SHEPPERD1997] M. Shepperd, and C. Schofield : Estimating software project effort using analogies, IEEE Trans. on Software Eng., Vol.23, No.12, pp.736-743, 1997
- [SRINIVASAN1995] K. Srinivasan, and D. Fisher : Machine learning approaches to estimating software development effort, IEEE Trans. on Software Eng., Vol.21, No.2, pp.126-137, 1995
- [TANAKA1995] 田中豊, 垂水共之 (編) : Windows版 統計解析ハンドブック 多変量解析, 共立出版, 東京, 1995
- [TSUNODA2005] 角田雅照, 大杉直樹, 門田暁人, 松本健一, 佐藤慎一 : 協調フィルタリングを用いたソフトウェア開発工数予測方法, 情報処理学会論文誌, Vol.46, No.5, pp.1156-1164, 2005
- [WALSTON1977] C. Walston, and C. Felix : A Method of programming measurement and estimation, IBM Systems Journal, Vol.16, No.1, pp.54-73, 1977
- [WINER1991] B. Winer, D. Brown, and K. Michels : Statistical Principles in Experimental Design ( 3rd edition ), McGraw-Hill, New York, 1991