

上流工程での活動実績を用いたソフトウェア開発工数見積もり方法の定量的評価

角田 雅照 戸田 航史 伏田 享平 亀井 靖高

Meiyappan Nagappan 鷓林 尚靖

本研究では、上流工程の活動実績に基づく工数見積もりに着目し、その見積もり精度を定量的に評価する。具体的には、上流工程での実績工数を説明変数とする開発総工数見積もりモデルを作成し、ソフトウェア規模に基づく工数見積もりモデルと精度を比較する。さらに、実績工数とソフトウェア規模の両方を説明変数とした場合のモデルとも比較する。実験では ISBSG データセットを用い、計画工程と要求分析工程を上流工程とみなしてモデルを構築した。実験の結果、計画・要求分析の合計工数とソフトウェア規模の両方を説明変数として用いることにより、見積もり精度が最も改善する (*BRE*(Balanced Relative Error) 平均値が 148.4%から 75.4%に改善する) ことがわかった。

We focus on effort estimation based on the effort for early phase activities, and built effort estimation models using early phase effort as an explanatory variable, and compared the estimation accuracies of these models to the effort estimation models based on software size. In addition, we built estimation models using both early phase effort and software size. In our experiment, we used ISBSG dataset, which was collected from software development companies, and regarded planning phase effort and requirement analysis effort as early phase effort. The result of the experiment showed that when both software size and sum of planning and requirement analysis phase effort were used as explanatory variables, the estimation accuracy was most improved (Average Balanced Relative Error was improved to 75.4% from 148.4%).

1 はじめに

近年、ソフトウェアはより大規模化しており、それにともないソフトウェア開発にはより多くの工数(コスト)が必要となっている。このような大規模なソフトウェア開発においては、スケジュールやコストに関するプロジェクト管理は必要不可欠であり、ソフトウェア開発工数見積もりは、その基礎となるものである。よって、工数見積もりには高い見積もり精度が求められる。ここで高い見積もりとは、見積もり工数と実績工数との差が小さいことを指す。主要な工数見積もり方法の1つとして、数学的モデルによる工数見積もりがあり、過去に蓄積された開発データに基づき、工数見積もりモデルを構築する。これまで、高精度な見積もりを実現するために多数のモデルが提案されている。

多くの工数見積もりモデルの研究では、ファンク

Empirical Evaluation of Software Development Effort Estimation Based on Upper Phase Development Activity.

Masateru Tsunoda, 近畿大学理工学部情報学科, Dept. of Informatics, Kinki University.

Koji Toda, 福岡工業大学情報工学科, Dept. of Information and Systems Engineering, Fukuoka Institute of Technology.

Kyohei Fushida, 株式会社 NTT データ技術開発本部, Research and Development Headquarters, NTT DATA Corporation.

Yasutaka Kamei, Naoyasu Ubayashi, 九州大学大学院システム情報科学研究院, Graduate School and Faculty of Information Science and Electrical Engineering, Kyushu University.

メイヤッパン ナガッパン, クイーンズ大学計数学科, School of Computing, Queen's University.

コンピュータソフトウェア, Vol.31, No.2 (2014), pp.129-143.

[研究論文] 2013 年 4 月 30 日受付.

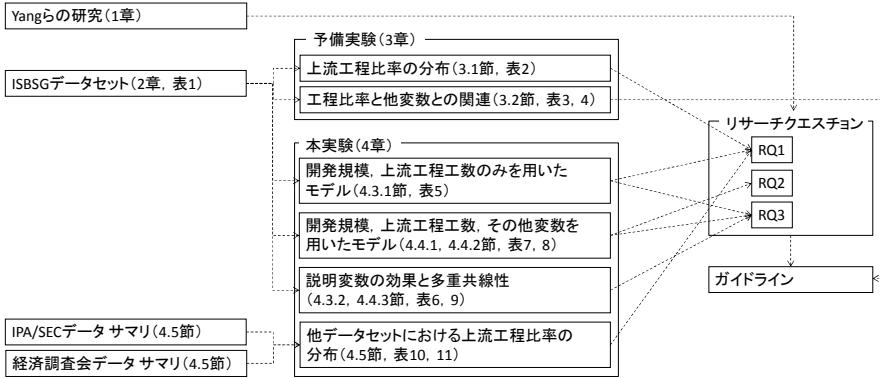


図 1 分析の概要

ションポイントなどのソフトウェアの規模(開発規模)を主要な説明変数(見積もり時に既知の変数)として開発工数を見積もっている。一方開発現場では、上流工程の工数と、その工程が開発工程全体に占める割合に基づいて、工数を簡易的に見積もる場合がある[5][18]。例えば、要件定義の開発工数全体に占める割合が過去のプロジェクトの平均で20%であり、ある開発中のプロジェクトの要件定義の工数が2人月である場合、総開発工数を10人月と見積もる。

本研究では、これまでの研究においてあまり重視されていなかった、要件定義などの上流工程の工数とその工程比率に基づいた工数見積もりに着目し、それらに基づく工数見積もりモデルの精度を評価する。ここで工程比率とは、当該工程にかけられた工数が開発工数全体に占める割合を指す。Yangらは開発工程の分布に関する分析を行った研究は非常に少ないことを指摘するとともに、上流工程(計画・要件定義、設計)の工程比率の分散はあまり大きくないことを示している[22]。このことから、上流工程にかけられた工数(以下、上流工程工数と呼ぶ)を説明変数として重回帰分析に基づく工数見積もりモデルを構築した場合、比較的高い見積もり精度となることが期待される(理由は3章で詳説する)。なお、工程比率はモデル構築時にパラメータとして推定されるため、説明変数に含まれない。

我々の知る限り、開発規模を用いたモデルと上流工程工数を用いたモデルの見積もり精度を比較した研究は存在せず、従って、見積もり精度の高いモデルを

構築するためには、どちらの変数を使うべきかは明らかでない。精度の差異を明らかにすることにより、どちらの変数を使うべきかを明らかにすることが本研究の目的である。そのために、以下の3つのリサーチ・クエスチョンを設定した。

- RQ1: 開発規模のみ、もしくは上流工程工数のみで工数を見積もった場合、どちらの見積もり精度が高いのか。
- RQ2: 開発規模、上流工程工数に加え、その他のデータ項目(開発言語など)を説明変数に加えた場合、開発規模を含むモデルと上流工程工数を含むモデルでは、どちらの見積もり精度が高いのか。
- RQ3: 開発規模と上流工程工数、両方を用いてモデルを構築した場合、見積もり精度はより高まるのか。そのときに多重共線性が発生しないのか(発生する場合、不安定なモデルとなるため開発規模と上流工程工数の両方を用いるべきではない)。

プロジェクトのデータ収集はコストが掛かるため、組織によってはデータを詳細に収集、蓄積していない場合がある。そこで、開発規模または上流工程工数のみを説明変数として用いることを想定した。RQ1を設定した。RQ1にその他の説明変数(開発言語など)を加えた場合、結論が変化するかを確かめるため、RQ2を設定した。RQ3はRQ1とRQ2両方の場合について確かめる。

リサーチ・クエスチョンに対する回答だけでは、精

表 1 分析に用いた変数

変数名	尺度	詳細
FP	比例	未調整 FP(調整係数の導出には問題点があり、かつ用いても精度が改善されない場合が多いため、未調整 FP の利用が好まれる [11])
総工数	比例	総開発工数
計画工数	比例	予備調査、プロジェクト計画作成などにかかった工数
要求分析工数	比例	要求分析、要件定義、アーキテクチャ定義などにかかった工数
計画・分析工数	比例	計画工数 + 要求分析工数
計画比率	比例	計画工数 / 総工数
計画・分析比率	比例	計画・分析工数 / 総工数
工期	比例	全開発期間 - プロジェクト停止期間
開発速度	比例	FP/工期
生産性	比例	総工数 / FP
開発種別	名義	プロジェクトの開発種別。カテゴリは新規開発、改修・保守開発、再開発
開発プラットフォーム	名義	開発したソフトウェアが稼働するプラットフォーム。カテゴリは MF(メインフレーム)、PC(パソコン)、MR(ミッドレンジ)
開発言語	名義	プロジェクトで用いた開発言語の種別。カテゴリは 2GL、3GL、4GL(nGL は第 n 世代言語を示す)、自動生成コード

度の高いモデルを構築するために、具体的に開発現場でどうすればよいか十分にわかりやすいとはいえない。そこで本論文では、研究結果の現場での適用を容易にするために、リサーチ・クエスチョンの回答に基づいた、見積もり精度の高いモデルを構築するためのガイドラインを示す。

以降の分析概要を図 1 に示す。図右側は論文で求めることを示し、リサーチ・クエスチョンとガイドラインが該当する。図中央は図右側を求めるための分析であり、図左側は論文で参照した既存研究とデータを示す。論文中で該当する箇所と図表をカッコ内に記す。リサーチ・クエスチョンの設定は Yang らの研究 [22] に基づいており、それぞれに対する回答は点線でつながれている分析により決定する。各手順で用いるデータは点線でつながれている。ガイドラインはリサーチ・クエスチョン及び予備実験「工程比率と他変数との関連」に基づいて決定する。

2 データセット

2.1 プロジェクトの抽出

分析に用いたデータセットは、ISBSG(International Software Benchmarking Standards Group) が 20ヶ国の組織から収集したものである [4]。用いたデータセットは Release 9 と呼ばれるバージョンであり、1989 年から 2004 年に実施されたソフトウェア開発プロジェクトが 3026 件、変数が 99 個含まれている。

データには欠損値が含まれている。欠損値とは値が記録されていないことを指す。分析に用いた変数を表 1 に示す。

分析対象のプロジェクトの条件を整えるため、Lo-can ら [12] が示した条件を満たしているプロジェクトを抽出した。条件は、具体的には「ファンクションポイント (FP) が IFPUG 法で計測されている」などである。次に計画工数、要求分析工数、設計工数、コーディング工数、テスト工数、インプリメント工数(ソフトウェアのリリース、インストールなどの工数)の合計と、総工数の誤差が 10%以内のプロジェクトを抽出した。誤差は、各工程工数合計 < 総工数の場合、 $1 - \frac{\text{各工程工数合計}}{\text{総工数}}$ 、それ以外の場合、 $1 - \frac{\text{総工数}}{\text{各工程工数合計}}$ とした。完全に誤差がないデータだけを抽出するとデータ数が減少するため、10%以内の誤差があるプロジェクトを含めて分析を行った。さらに FP が計測されているプロジェクトのみを抽出した。

予備分析では、上記の条件で抽出したプロジェクト 172 件を用いた。RQ1 に関する分析では、それらのプロジェクトから、計画工数と要求定義工数の両方が計測されているプロジェクト 118 件を抽出した。RQ2 に関する分析では、それらから開発規模、開発プラットフォーム、開発言語が記録されている 70 件のプロジェクトを抽出した。

2.2 上流工程工数の定義

上流工程工数は説明変数として用いるため、上流工程工数と見なすことができるのは、見積もりモデル利用時(総工数見積もり時期)までに確定している工程の工数である。本研究では基本設計完了後をモデル利用時と想定し、基本設計工数までを上流工程工数に含むことができるとした。ただし ISBSG データでは基本設計と詳細設計の工数が「設計工数」として記録されているため、分析では基本設計工数は含めずに要求分析工数までを上流工程工数とした。なお、基本設計完了後をモデル利用時と想定した理由は、分析対象のプロジェクトにおいて開発規模が確定したと考えられる時期であるためである。分析対象のプロジェクトは IFPUG 法で計測されており、該当手法で規模を計測するには基本設計が完了している必要がある。

以降、上流工程工数とは、ISBSG データに関する箇所については計画工程までの工数と分析工程までの工数を指すが、それ以外の箇所では最大で基本設計工程までの工数を指す。なお、分析工程までの工数を説明変数に用いるよりも、定義の最大値である、基本設計工程までの工数を用いるほうが、見積もり精度の面で有利であると考えられる。これは、前者では基本設計、詳細設計、製造、試験の4つの工程の工数が未確定であるが、後者では3つの工程の工数が未確定であり、かつ残り1つの工程の工数は説明変数として使えるためである。

計画工数と要求分析工数の相関係数は0.66と大きくっており、別々に説明変数とすると多重共線性が発生する可能性が高まる。そこで、多重共線性を避けるため、見積もりモデル構築時には計画工数と要求分析工数それぞれを説明変数として用いずに、それらを合計した計画・分析工数を説明変数として用いる。相関が大きい理由は、開発規模が大きくなると、計画立案や要求分析に手間が掛かり、どちらの工程も工数がより多く必要となるためである。超上流で見積もりを行う場合を想定し、計画工数を説明変数として用いた場合の見積もり精度についても確かめる。

3 予備分析

3章では2つの予備分析を行う。3.1節ではRQ1に関する分析の信頼性を高めることを目的として、生産性と上流工程比率の分布を比較する。3.2節ではガイドラインに反映させることを目的として、上流工程工数を用いたモデルに加えるべき説明変数を検討する。

3.1 上流工程比率の分布

数学的モデルによる工数見積もりでは、自社内や企業横断的に収集されたデータを用いて見積もりモデルを構築する[12]。モデル構築方法として対数(重)回帰分析がよく用いられる[17]。これは、工数や開発規模を対数変換してから(重)回帰分析を行う手法であり、説明変数を1つだけ用いた最も単純なモデルは以下の式となる。

$$y = ax^b \quad (1)$$

ここで y は見積もり工数を表す。対数(重)回帰分析とは、工数や開発規模を対数変換してから(重)回帰分析を行う手法である。 b は定数であり、一般に1に近い値となる。開発規模を説明変数として用いる場合、 x は開発規模であるため、 a は生産性(の逆数)と見なせる。大まかには a は過去プロジェクトの生産性の平均値となり、モデル構築時に推定される。説明変数が1つだけだと a は定数となるため、生産性の分散、すなわちプロジェクトによる差異が大きき場合、モデルがプロジェクトにあてはまりにくくなる。

同様に上流工程工数を用いる場合、 x は上流工程工数であるため、 a は上流工程の工程比率(の逆数)と見なせる。大まかには a は過去プロジェクトの上流工程比率の平均値となる。分散が小さい場合、モデルがプロジェクトにあてはまりやすくなり、精度が高まることが期待される。

そこで RQ1 に関する分析の信頼性を高めるために、生産性と上流工程比率の分布を比較し、見積もり精度の差異が発生する原因を把握する(生産性と上流工程比率の定義は表1参照)。なおここでは、開発規模や上流工程工数以外の説明変数、例えば開発言語などのプロダクトの特性を示す変数を用いずに見積もりモデ

表 2 計画比率, 計画・分析比率, 生産性の分布

	平均値	最小値	Q ₁	中央値	Q ₃	最大値	Q ₃ /Q ₁	最大値 / 最小値
計画比率	0.09	0.005	0.039	0.071	0.124	0.434	3.2	86.8
計画・分析比率	0.223	0.018	0.133	0.208	0.294	0.614	2.2	33.3
生産性	0.197	0.003	0.057	0.127	0.288	1.044	5.1	348

ルを構築することが通常であると想定しているのではない。1章でも述べたように、プロジェクトのデータ収集はコストが掛かるため、組織によってはプロダクトの特性を示す変数を十分に収集、蓄積していない場合がある。例えば、文献[19]中で示されているデータセットでは、用いることができる説明変数は開発規模のみとなっている。そのため、プロダクトの特性を示す変数などがやむなく使えない場合を想定している。

表2に計画比率, 計画・分析比率, 生産性の分布を示す。分析では分散を用いずに第1四分位数と第3四分位数の比(Q₃/Q₁)と最大値と最小値の比(最大値 / 最小値)(表中太字)を用いた。前者は外れ値による影響を除外するため、後者は逆に外れ値の影響に着目するために用いた。外れ値とは他のケースと比べて、値が極端に大きい、または小さいケースを指す。Q₃/Q₁、及び最大値 / 最小値を比較すると、計画比率と計画・分析比率よりも生産性のほうが大きかった。すなわち、計画比率, 計画・分析比率よりも、生産性のほうがプロジェクト間の差異が大きいといえる。これは、RQ1の回答が「上流工程工数のみで工数を見積もった場合、開発規模のみで工数を見積もった場合よりも精度が高くなる」となることを示唆している。

3.2 工程比率と他変数との関連

上流工程の工程比率と関連の強い変数は、説明変数の候補として採用する、すなわちモデル構築のガイドラインに、それらを必要性の高い説明変数として記述する必要がある。単変量解析、ここでは分散分析の寄与率とSpearmanの順位相関係数を適用し、工程比率とその他の変数との関連を明らかにする。

また、説明変数として用いることができない変数(総工数, 生産性, 工期, 開発速度)についても、工程比率との関連を分析した。これらの変数が工程比率との関連が強い場合、上流工程工数を用いたモデルを適用する場合に注意が必要となる。例えば、もし開発

速度と工程比率との関連が強く、かつ開発速度がプロジェクトによって大きく異なる場合、見積もり精度が低くなると考えられるため、モデルを用いるべきではない。よって、モデルの信頼性を高めるためには、これらの関連を分析する必要がある。

なお、開発プラットフォームと開発言語は上流工程において用いないため、これらは上流工程比率と無関係に見える。ただし、以下の例のように開発言語が上流工程以外の工数と関連がある場合、上流工程比率とも関連が生じうる。

- プロジェクト A: 計画・分析工数が100人時、その他の工数が300人時であり、計画・分析比率は25%である。
- プロジェクト B: 計画・分析工数が100人時、その他の工数が400人時であり、計画・分析比率は20%である。
- その他の工数の差の原因は開発言語である。

この場合、プロジェクト間で計画・要件定義「工数」は同じだが、計画・要件定義「比率」は異なる。同様の理由で、上流工程の「工数」との関連が強く、上流工程の「比率」との関連が弱い変数も存在しうるが、分析目的と直接関連しないため、ここでは分析対象としない。

まず、名義尺度の変数と工程比率との関係について、分散分析の調整済み寄与率 ω^2 [21]を用いて分析した。名義尺度とは、開発言語などの数値で表されない変数を指す。 ω^2 は名義尺度の変数と比例尺度の変数の関連の強さを分析する際に用いられ、おおむね0から1の値をとり、値が大きいほど関連が強いことを示す。工程比率と名義尺度変数との寄与率と、分散分析のp値を表3に示す。*は有意水準5%で関連があったものを表す。p値が有意水準を下回る場合、統計的に関連があるとみなされる。開発プラットフォームは計画・分析比率と有意な関連があったが、あまり

表 3 工程比率と名義尺度変数との関連

変数		開発種別	開発プラットフォーム	開発言語
計画比率	ω^2	0.01	0.03	-0.01
	p 値	0.18	0.1	0.52
計画・分析比率	ω^2	-0.01	0.08*	-0.01
	p 値	0.74	0.01	0.5

表 4 工程比率と比例尺度変数との関連

		FP	総工数	生産性	工期	開発速度
計画比率	ρ	-0.1	-0.17	0.12	-0.28*	0.01
	p 値	0.25	0.05	0.18	0	0.88
計画・分析比率	ρ	-0.20*	-0.20*	0.06	-0.18	-0.05
	p 値	0.03	0.03	0.51	0.08	0.62

強い関連ではなかった。その他の変数と工程比率は関連が見られなかった。

次に、比例尺度の変数と工程比率との関係について、Spearman の順位相関係数 ρ を用いて分析した。比例尺度とは、数値で表され加減乗除できる変数を指す。Spearman の順位相関係数は -1 から 1 の値をとり、絶対値が 1 に近いほど変数間の関連が強いことを示す。ソフトウェア開発のデータセットでは値の大きいケース（開発規模の大きなケースなど）が含まれるなど、正規分布を仮定できないことが多く、Pearson の積率相関係数では正しく関係を分析できない場合があるため用いていない。工程比率と比例尺度変数との相関係数と p 値を表 4 に示す。* は有意水準 5% で関連があったものを表す。工期と計画比率、FP と計画・分析比率、総工数と計画・分析比率に有意な関連が見られたが、それぞれあまり強い関連ではなかった。その他の変数と工程比率には関連が見られなかった。

データセットに含まれる変数に対して単変量解析を適用した限りにおいては、名義尺度、比例尺度の変数どちらも工程比率との関連が強くなかった。また、説明変数として用いることができない変数（総工数、生産性、工期、開発速度）についても、工程比率との関連は強くなかった。すなわち、分析では上流工程工数を用いた見積りモデルの説明変数の候補に必ず含めるべき（モデル構築のガイドラインに、必要性の高い説明変数として記述すべき）変数は見つからなかった。

4 上流工程工数を用いた工数見積り

4.1 実験概要

RQ1, RQ3 に回答するために、118 件のプロジェクトを用いて 5 種類の工数見積りモデルを構築し、見積り精度を比較した。5 種類のモデルそれぞれの説明変数と、対応するリサーチクエスチョンを以下に示す。

- モデル I: FP(RQ1)
- モデル II: 計画工数 (RQ1)
- モデル III: 計画・分析工数 (RQ1)
- モデル IV: 計画工数, FP(RQ3)
- モデル V: 計画・分析工数, FP(RQ3)

RQ2, RQ3 に回答するために、上記の変数に開発種別、開発プラットフォーム、開発言語を説明変数の候補として加え、70 件のプロジェクトを用いてモデルを構築した。この場合、前出の箇条書きの RQ1 の箇所は RQ2 となる。これらの変数は Locan ら [12] が用いた変数と同じである。予備分析の結果からは、これらの変数は上流工程工数を用いた見積りモデルの精度向上に大きく寄与しないと考えられるが、精度向上に全く役立たないとも限らない。そこで説明変数の候補として用いた。

比例尺度の変数を対数変換し、回帰分析を用いて見積りモデルを構築した。モデルに多重共線性が発生しているかを確かめるために VIF (分散拡大要因) を用いた。各変数の VIF が 10 を超える場合、多重共線性が発生しているとされる [20]。

モデル構築時には、5-fold cross validation を用いてモデル構築用のフィットデータとモデル評価用のテストデータにデータセットを分割した。5-fold cross validation とは、あるデータを用いて構築されたモデルが、別のデータでも有効に働くかを確かめるための方法である。具体的には、(1) データをランダムに 5 分割する、(2) 1/5 のデータをテストデータ、残り 4/5 をフィットデータとする、(3) テストデータとフィットデータの組み合わせを変更し、(2) を合計 5 回繰り返す。試行回数が少ないと統計的にモデルの見積り精度差を確かめることが難しいため、5-fold cross validation を 4 回繰り返して試行回数を増やし

た(合計 20 回の試行となる)。

RQ2, RQ3 に対応するための, 開発言語などを説明変数の候補に含むモデルを構築する時には, AIC(赤池情報量規準)に基づくステップワイズ変数選択を行った。変数選択は, 見積もりモデルから不要な説明変数を除去することにより, より適切なモデルを得るための方法である。

名義尺度の変数はダミー変数化した。ダミー変数化とは, 例えば開発種別が新規開発の場合は 1, それ以外の場合は 0 の値をとるような変数を作成することであり, これにより, 名義尺度の変数を重回帰分析などで利用できるようになる。変数名は表 1 の各カテゴリ名を用いた。なお, 開発種別が再開発, 開発言語が 2GL または自動生成コードについては, それぞれプロジェクトが 1 件だったため, それぞれを示すダミー変数を削除した。

説明変数を 2 つ以上用いた場合, モデル I から III のモデル式は以下ようになる。

$$y = e^{b_0 + b_1 x_1 + b_2 x_2} x_3^{b_3} \quad (2)$$

ここで y は見積もり工数を表し, e は自然対数の底, b_0 は切片, b_1 から b_3 は回帰係数を表す。 x_1, x_2 は開発言語などのカテゴリ変数を表し, x_3 は開発規模, または上流工程工数を表す。モデル IV, V のモデル式は以下ようになる。

$$y = e^{b_0 + b_1 x_1 + b_2 x_2} x_3^{b_3} x_4^{b_4} \quad (3)$$

この場合, x_3 は開発規模, x_4 は上流工程工数を表す。

4.2 評価指標

工数見積もり精度の評価指標として, AE (Absolute Error), MRE (Magnitude of Relative Error) [2], MER (Magnitude of Error Relative to the estimate) [9], BRE (Balanced Relative Error) [15] の 4 つの指標の平均値を用いた。工数の実測値を x , 見積もり値を \hat{x} とするとき, それぞれの指標は以下の式により求められる。

$$AE = |x - \hat{x}| \quad (4)$$

$$MRE = \frac{|x - \hat{x}|}{x} \quad (5)$$

$$MER = \frac{|x - \hat{x}|}{\hat{x}} \quad (6)$$

$$BRE = \begin{cases} \frac{|\hat{x} - x|}{x}, & \hat{x} - x \geq 0 \\ \frac{|\hat{x} - x|}{\hat{x}}, & \hat{x} - x < 0 \end{cases} \quad (7)$$

以降, 例えば AE の平均値を MAE などと表し, 中央値を $MdAE$ などと表す。各指標の値が小さいほどモデルの精度が高いことを示す。ただし, MRE と MER は過大見積もりと過小見積もりに対し, アンバランスな評価になるという問題がある [1][11]。そこで本研究では, それらをバランスよく評価する指標 [16] である BRE を特に重視してモデルの精度を評価した。さらに BRE 以外の評価指標が改善しているかを確かめることにより, 実験結果の信頼性を高める。もし BRE 以外の多くの指標が改善していない場合, そのモデルの見積もり精度は必ずしも高いとはいえない。

モデルの見積もり精度は, ベースラインとしたモデルとの評価指標の差により評価した。すなわち, 値が大きいほどベースラインよりも精度が改善したことを示し, 負となっている場合, ベースラインよりも精度が悪化したことを示す。差の検定にはノンパラメトリックな検定である Wilcoxon の符号付順位和検定を用い, 有意水準は 5% とした。差の検定とは, 2 つの評価指標の差が統計学的に確かであることを調べることであり, ノンパラメトリックな検定は外れ値の影響を受けにくいという特長がある。

4.3 開発規模, 上流工程工数のみを用いたモデル

4.3.1 見積もり精度の比較

見積もり精度を表 5 に示す。表の 2 行目はモデル I (ベースライン) の見積もり精度であり, 3 行目以降はモデル I (2 行目) との見積もり精度の差分 (20 回の試行の平均値) を示している。負の値はモデル I (FP を用いたモデル) よりも精度が悪化したことを示す。* は有意差があった場合を示している。見積もり精度に関して以下の結果が得られた。

- モデル II, III は, 全ての評価指標においてモデル I よりも改善しており, モデル II の $MdAE$ と $MdMRE$ を除いて差が有意であった。すなわち, 上流工程工数を用いると, 開発規模を用いるよりも精度が高くなるといえる。

表 5 FP, 上流工程工数のみを用いたモデルの見積もり精度

モデル		MAE	MdAE	MMRE	MdMRE	MMER	MdMER	MBRE	MdBRE
I(FP)		3725.6	1489.9	102.90%	60.30%	113.40%	62.70%	168.80%	104.10%
II(計画工数)	差分	373.3*	122.3	28.2%*	5.60%	28.3%*	3.8%*	52.5%*	22.4%*
	p 値	0	0.28	0	0.12	0.03	0	0	0.01
III(計画・分析工数)	差分	1214.1*	437.8*	49.5%*	17.6%*	53.5%*	20.0%*	91.6%*	45.1%*
	p 値	0	0	0	0	0	0	0	0
IV(計画工数, FP)	差分	805.2*	365.5*	37.1%*	8.6%*	40.3%*	17.4%*	69.8%*	33.7%*
	p 値	0	0	0	0	0	0	0	0
V(計画・分析工数, FP)	差分	1512.4*	681.3*	57.6%*	26.3%*	61.7%*	31.1%*	103.1%*	64.7%*
	p 値	0	0	0	0	0	0	0	0

表 6 FP と上流工程工数を用いたモデルの偏回帰係数

(a) モデル I			(b) モデル II			(c) モデル III		
変数	β	VIF	変数	β	VIF	変数	β	VIF
FP	0.63	1	計画工数	0.77	1	計画・分析工数	0.87	1
(d) モデル IV			(e) モデル V					
変数	β	VIF	変数	β	VIF			
FP	0.33	1.34	FP	0.25	1.37			
計画工数	0.61	1.34	計画・分析工数	0.75	1.37			

- モデル IV, V の見積もり精度は, モデル II, III よりも高かった. また, モデル I と比較すると大きく精度が改善しており, 全ての評価指標で有意差があった. すなわち, 上流工程工数と開発規模の両方を用いると見積もり精度が改善するといえる.
- モデル V の見積もり精度が最も高く, モデル III の見積もり精度はモデル II, IV よりも高かった. よって, 要求分析工数を用いると, 精度が大きく改善するといえる.

4.3.2 説明変数の効果と多重共線性

計画工数, 計画・分析工数, FP を用いたモデルの標準化偏回帰係数 (β) と VIF を表 6 に示す (20 回の試行の平均値). 標準化偏回帰係数が大きいほど, 目的変数, すなわち見積もり工数に対する影響が大きいことを示す. 構築されたモデルについて, 以下の結果が得られた.

- モデル IV, V (FP と上流工程工数両方を用いたモデル) では, どの試行においても VIF が 10 を超えることはなかった. よって, 開発規模と上流工程工数両方を説明変数として用いても, 多重共

線性は発生しないといえる.

- モデル IV, V では, 上流工程工数の偏回帰係数のほうが開発規模よりも大きかった. これは, 上流工程工数の見積り工数に対する影響が, 開発規模よりも大きいことを示している. すなわち, 上流工程工数はモデルに不可欠であるといえる.

4.3.3 リサーチ・クエスチョンに対する回答

分析結果より, RQ1 に対する回答は「(開発言語などの他の変数を用いずに) 開発規模のみ, もしくは上流工程工数のみで工数を見積もった場合, 後者の見積もり精度が高い」となる. (RQ2 を考慮しない場合の) RQ3 に対する回答は「開発規模と上流工程工数両方を説明変数として用いると, 工数見積もり精度が高まる. 多重共線性は発生しない」となる.

4.4 開発規模, 上流工程工数, その他変数を用いたモデル

4.4.1 その他の説明変数追加の効果

その他の変数 (開発種別, プラットフォーム種別, 開発言語) を説明変数に加えた場合の効果を確認する

表 7 FP, その他の変数 (開発種別, 開発プラットフォーム, 開発言語) を用いたモデルと用いないモデルの見積もり精度

モデル		MAE	MdAE	MMRE	MdMRE	MMER	MdMER	MBRE	MdBRE
I(FP, その他変数あり/なし)	差分	-55.8	-15	22.7%*	0.60%	-4.30%	4.70%	16.20%	10.20%
	p 値	0.99	0.84	0	0.7	0.87	0.12	0.09	0.13
II(計画工数, その他変数あり/なし)	差分	-112.8	-10.4	-2.60%	-3.60%	-8.1%*	-1.50%	-9.70%	-6.5%*
	p 値	0.01	0.72	0.34	0.09	0	0.34	0.05	0.02
III(計画・分析工数, その他変数あり/なし)	差分	-122.9	-4.6	-1.80%	-2.30%	-4.6%*	0.40%	-5.30%	-0.20%
	p 値	0.52	0.52	0.65	0.59	0.02	0.65	0.05	0.78
IV(計画工数, FP, その他変数あり/なし)	差分	-212.4*	37.9	1.60%	0.00%	-3.50%	-2.90%	-2.10%	-3.10%
	p 値	0.02	0.23	0.43	0.99	0.9	0.26	0.78	0.62
V(計画・分析工数, FP, その他変数あり/なし)	差分	-166.4*	-124	-0.90%	-0.90%	-1.20%	-1.70%	-1.70%	-2.90%
	p 値	0.01	0.06	0.69	0.61	0.51	0.18	0.46	0.41

表 8 FP とその他の変数を用いたモデル, FP と上流工程工数を用いたモデルの見積もり精度

モデル		MAE	MdAE	MMRE	MdMRE	MMER	MdMER	MBRE	MdBRE
I(FP, その他変数)		2966.6	1202.3	76.50%	48.70%	112.80%	48.80%	148.40%	72.50%
II(計画工数)	差分	-87.4	-81.9	-0.80%	-8.0%*	21.10%	-0.30%	22.90%	-7.40%
	p 値	0.55	0.39	0.93	0.04	0.62	0.81	0.81	0.28
III(計画・分析工数)	差分	230.5	245.4*	23.9%*	0.80%	42.90%	8.8%*	62.0%*	11.80%
	p 値	0.15	0.04	0	0.6	0.05	0.01	0	0.07
IV(計画工数, FP)	差分	234.4*	193.9*	14.0%*	2.60%	37.80%	8.9%*	48.9%*	14.5%*
	p 値	0.04	0.01	0	0.57	0.11	0.02	0.01	0.1
V(計画・分析工数, FP)	差分	540.0*	527.9*	30.2%*	11.2%*	51.8%*	16.2%*	73.1%*	28.8%*
	p 値	0	0	0	0.01	0	0	0	0

表 9 FP, 上流工程工数, その他の変数 (開発種別, 開発プラットフォーム, 開発言語) を用いたモデルの偏帰係数

(a) モデル I				(b) モデル II				(c) モデル III			
変数	β	VIF	採用回数	変数	β	VIF	採用回数	変数	β	VIF	採用回数
FP	0.71	1.07	20	計画工数	0.74	1.08	20	計画・分析工数	0.87	1.11	20
3GL	0.27	1.14	20	3GL	0.17	1.05	12	3GL	0.13	1.11	8
MF	0.24	1.07	20	MF	-0.18	1.2	8	MF	-0.16	1.17	16
MR	0.14	1.82	1					MR	-0.12	1.52	1
新規開発	0.16	1.41	1								
(d) モデル IV				(e) モデル V							
変数	β	VIF	採用回数	変数	β	VIF	採用回数				
FP	0.44	1.33	20	計画・分析工数	0.66	1.6	20				
計画工数	0.5	1.4	20	FP	0.3	1.52	20				
3GL	0.19	1.06	20	3GL	0.14	1.11	19				
MF	0.18	1.52	3	MF	0.14	1.75	1				
				新規開発	-0.09	1.25	2				

ために、これらの変数を用いたモデルと用いないモデル同士を比較した。これらの変数を用いていないモデル (4.3 節で構築したモデル) をベースラインモデルとし、見積もり精度を比較した結果を表 7 に示す。

表はモデル I から V それぞれにおいて、その他の

変数を用いないモデルと、その他の変数を用いたモデルとの差分を示す。表 5 のモデル I は差分ではなく実際の評価指標の値を示すが、表 7 ではモデル I も差分を表すことに注意すること。表 7 において、負の値はこれらの変数を用いることにより精度が悪化

したことを示す。モデル I は精度が改善したが、その他のモデルは精度がわずかに悪化していた。従って、これらの変数を説明変数に加えても、上流工程工数を用いたモデルの見積もり精度は改善しないといえる。

4.4.2 見積もり精度の比較

表7の結果に基づき、ベースラインモデルはFPとその他の変数を用いたモデル I とし、モデル II から V ではその他の変数を用いないモデルを比較対象とした。各モデルの見積もり精度を表8に示す。見積もり精度に関して以下の結果が得られた。

- モデル II の評価指標は、8 個中 6 個がモデル I より悪化していた。すなわち、計画工数を用いたモデルは開発規模を用いたモデルよりも見積もり精度が低いといえる。
- モデル III では、全ての評価指標がモデル I よりも改善しており、*MBRE* については有意に改善していた。よって、計画・分析工数を用いたモデルは開発工数を用いたモデルよりも見積もり精度が高いといえる。
- モデル IV, V それぞれは、モデル II, III よりも見積もり精度が高かった。また、モデル IV, V はモデル I よりも見積もり精度が改善しており、*MBRE* については有意に改善していた。従って、開発規模と上流工程工数の両方を用いることにより、見積もり精度が改善できるといえる。

4.4.3 説明変数の効果と多重共線性

FP, 上流工程工数とその他の変数を用いたモデルの標準化偏回帰係数と VIF を表9に示す。採用回数とは、20 回の試行のうち説明変数として採用された回数を示す。構築されたモデルにおいて、以下の結果が得られた。

- どの試行においても VIF が 10 を超えることはなかった。
- FP と上流工程工数の偏回帰係数の大きさを比較すると、モデル V では計画・分析工数のほうが FP よりも係数が大きかったが、モデル IV では計画工数と FP では差が小さかった。
- その他の変数では、3GL と MF が説明変数として採用される回数が多かったが、モデル IV, V では MF の採用回数は低かった。開発種別はど

のモデルでもほとんど採用されなかった。

4.4.4 リサーチ・クエスチョンに対する回答

分析結果より、RQ2 に対する回答は「開発規模を用いたモデルと比較して、計画工数を用いたモデルは見積り精度が低く、計画・分析工数を用いたモデルは見積もり精度が高い」となる。また、前節の分析結果と合わせ、RQ3 に対する回答は「開発規模と上流工程工数、両方を用いてモデルを構築した場合、見積もり精度はより高まる。その他の変数の有無に関わらず、多重共線性は発生しない」となる。

4.5 他データセットにおける上流工程比率の分布

3章での予備分析と同様に、他のデータセットにおける上流工程比率と生産性の分布を比較することにより、分析結果の信頼性を高める。3章と同様に、ここではプロダクトの特性などを示す説明変数がやむなく使えない場合を想定している。文献[19]のデータセットの他に、文献[10]のデータセットにおいても見積もりに利用できる変数が非常に少なく、開発規模以外では開発種別しか存在しない。後述するように分析では開発種別で層別しており、開発種別は考慮して分布を比較しているといえる。

比較対象は、経済調査会と情報処理推進機構・ソフトウェアエンジニアリングセンター (IPA/SEC) それぞれが収集したデータセットである。Yang ら [22] が用いている CSBSG (China Software Benchmarking Standard Group) データセットは、一部の統計量が不明であるため比較対象としなかった。経済調査会のデータセット [8] は、2001 年から 2008 年にかけて、日本の情報処理サービス企業 268 社から収集されたものであり、中小企業から大企業まで幅広い層から収集していることに特色がある。IPA/SEC のデータセット [6] は、大手 IT ベンダー 23 社から 2000 年代のプロジェクトを収集したものである。

どちらのデータセットにおいても、計画工程比率と要求分析工程比率は示されていなかった。そこで、基本設計工程の比率を用いて分析した。基本設計工程の比率は、基本設計工数 ÷ 基本設計からシステムテストまでの工数合計により計算される (以降基本設計比率とする)。統計量は新規開発と改造開発に分けて示さ

表 10 経済調査会データにおける基本設計比率, 生産性の分布 [8]

		平均値	最小値	Q ₁	中央値	Q ₃	最大値	Q ₃ /Q ₁	最大値 / 最小値
新規	基本設計比率	0.157	0.03	0.1	0.15	0.2	0.66	2	22
	生産性	20.6	0.4	8.2	15.7	23.2	189.5	2.8	473.8
改造	設計比率	0.141	0.029	0.1	0.138	0.2	0.28	2	9.7
	生産性	23.6	3.3	7.7	15.2	23.8	118.5	3.1	35.9

表 11 IPA/SEC データにおける基本設計比率, 生産性の分布 [6]

		平均値	最小値	Q ₁	中央値	Q ₃	最大値	Q ₃ /Q ₁	最大値 / 最小値
新規	基本設計比率	0.161	0.001	0.095	0.143	0.205	0.589	2.2	589
	生産性	18	0.8	7.7	11.9	20.1	118.2	2.6	140.7
改造	設計比率	0.147	0.002	0.095	0.137	0.189	0.557	2	278.5
	生産性	26	0.3	7.8	14.8	34.9	235.8	4.5	873.2

れていた。経済調査会データでのFPの計測手法は主にIFPUG法(一部例外あり)、IPA/SECデータではIFPUG法のみである。工数の計測単位は人月である。

経済調査会データにおける基本設計比率, 生産性の分布を表10に, IPA/SECデータにおける分布を表11に示す。IPA/SECデータの新規開発での最大値 / 最小を除き, Q₃/Q₁及び最大値 / 最小値(表中太字)は, 基本設計比率に比べて生産性のほうが差異が大きかった。このことから, 他のデータセットにおいても上流工程比率, または開発規模以外の説明変数を使わない場合, 上流工程比率を用いたモデルの見積もり精度が高くなる, すなわちRQ1に対する回答は他のデータセットにおいても正しい, と考えられる。

4.6 工数見積もりモデル構築のガイドライン

RQ1~RQ3への回答に基づいた, 工数見積もりモデル構築の新たなガイドラインを以下に示す。

- (4.3.1の分析結果より) 収集しているデータ項目が非常に少ない組織では, 上流工程工数のみで総工数を見積もるモデルを作成するとよい。このモデルは比較的高い見積もり精度が期待できる。
- (4.3.1, 4.3.2, 4.4.2, 4.4.3の分析結果より) 開発規模をファンクションポイント法などにより正確に計測している組織では, 開発規模に加えて上流工程工数を工数見積もりモデルに用いるとよい。多重共線性が起きる可能性が低く, また精度向上が期待できる。
- (4.4.1の分析結果より) 項目を多数収集してい

る組織において, 上流工程工数を用いたモデルを構築する場合, それらの項目を安易に説明変数に用いるべきではない。それにより見積もり精度が低下する可能性がある。もし用いるなら, 事前に本論文の実験と同様の方法でモデルを構築して見積もり精度を評価すべきである。ただし, 同様の実験は手間がかかるため, 実施が難しい場合, それらの項目を用いずにモデルを構築するとよい。

5 考察

5.1 上流工程工数の妥当性

ISBSGの定義では, 設計工程は基本設計, 詳細設計などの工数を示すが, 設計工数が記録されたプロジェクトは172件中4件のみであった。データ提供組織がこれらの工数を要求分析工数または次工程のコーディング工数に含めた可能性がある。すなわち, 計画・分析工数は基本設計工数を含んでいる可能性が高く, 場合によっては, 上流工程工数と見なされない詳細設計工数まで含んでいることも考えられる。その場合, 表5, 表8におけるモデルIII, Vの見積もり精度は, 上流工程工数を用いたモデルの精度を正しく表していないことになる。

計画・分析工数が詳細設計工数を含んでいるかどうかを確かめるために, Yangらの分析結果[22]を用いてCSBSGデータセットとの比較を行った。CSBSGデータセットでは, 計画・分析比率は16.1%, 計画・分析比率と設計比率の合計は31.0%であった。ISBSGデータでは計画・分析比率は22.3%であることから,

計画・分析工数に基本設計工数は含まれているが、詳細設計工数はあまり含まれていないと推測される。

5.2 手戻り作業が上流工程工数に及ぼす影響

モデルの見積もり精度を高めるためには上流工程工数を正確に計測することが重要である。ただし、ソフトウェア開発では手戻り作業が発生するため各工程の区切りが曖昧となりやすい。そのため、上流工程工数を正確に計測することは容易ではない。上流工程工数を正確に計測する代わりに、手戻り作業を考慮せずにある時点までの工数を計測し、それを上流工程工数と見なせば、計測の基準が統一され、見積もり精度が高まることが期待される。例えば、委託契約から請負契約への変更時までにかかった工数を上流工程工数と見なすことなどが考えられる。

上述のように、手戻り作業が発生して各工程の区切りが曖昧となりやすいため、上流工程工数には誤差が含まれる可能性がある。ただし、目的変数である総工数については、プロジェクトの終了(解散)時点が明確であるため上記要因を考慮する必要がなく、上流工程工数と比べて誤差が小さいと考えられる。すなわち、上流工程工数の誤差はモデルの見積もり精度に悪影響を与えている可能性があるが、見積もり精度の評価の妥当性には影響を与えていないと考えられる。

上流工程工数を用いた見積もりモデルは、上流工程工数が小さい場合、総工数も小さくなることを想定している。実際には、上流工程の作業が不十分のために上流工程工数が小さく、それ以降の工程の工数が増大しているプロジェクトが存在する。このプロジェクトにはモデルの前提が一致しないため、工数見積もり精度が低下する。ただし、実験では上流工程工数を用いた見積もりモデルの精度が高かったことから、そのようなプロジェクトは比較的少ない、すなわち開発規模を用いたモデルにおいて、規模が小さいにもかかわらず工数が大きくなることのほうが多いと考えられる。

5.3 開発規模の計測

基本設計よりも前に開発規模を見積もる手法がいくつか存在する(NESMA法[3]など)。これらの手法

を用いて基本設計よりも前に開発規模を見積もり、それに基づき工数を見積もることを前提とした場合、RQ1、RQ2に対する回答が異なる、すなわち、モデルIとIIIの比較結果より、開発規模を用いたほうが精度が高くなる可能性がある。ただしその場合でも、RQ3に対する回答は有用であると考えられる。要求分析をせずに開発規模を正確に見積もることは困難であり、要求分析にはある程度の工数を要することから、開発規模を見積もった時点で、ある程度の工数がかかっている。RQ3に関する分析結果より、その工数と開発規模を用いれば、総工数見積もりの精度が高まると期待される。

分析結果は、開発規模の計測、見積もりをしない場合でも、ある程度の精度で工数を見積もることができるとを示している。ソースコード作成前に開発規模を計測、見積もることは非常に容易とまではいえないため、開発企業によっては工数見積もり時に開発規模が正確に決定していない可能性がある。分析結果はそれらの企業にとって有用であると考えられる。

5.4 見積もり誤差の影響の大きさ

モデルIとモデルII~Vとの精度の現実での差は、見積もり対象のプロジェクトの開発費の大きさによって変化する。工数はおおむね開発費とみなすことができる。情報処理推進機構[7]の調査によると、約4割のプロジェクトの開発費が1000万円以上となっている。表5の*MdBRE*より、モデルIの*BRE*を104.1%、モデルVの*BRE*を39.4%とすると、1000万円のプロジェクトを見積もった場合の両者の見積もりの差は227万円となる。この差は開発費が大きくなるほど顕著となり、例えば2000万円のプロジェクトでの差は455万円となる。

6 関連研究

丸地ら[14]は、各工程の工数と、各工程における手戻り工数と不具合数を用いて、不具合削減施策を行って不具合数が減った場合、どの程度工数が減るのかを見積もる(評価する)モデルを提案している。ただし、このモデルの目的は不具合削減施策の効果を評価することであり、各工程の工数は既知であることが前提

となっているため、一般的な、プロジェクト計画立案のための工数見積りに利用することはできない。

上流工程の工数と、その工数が全工数に占める割合に基づいて工数を簡易的に見積もる方法は、いくつかの文献で紹介されている[5][18]。文献[5]では、プロジェクトの進捗を確認するために、あるフェーズ終了後の実績工数と、その工数が全工程に占める割合の平均値を用いて工数を再見積りする方法が紹介されている。文献[18]では、要求分析工数を見積り、その見積り工数と要件定義が全工程に占める割合から、工数を見積もる方法を紹介している。MacDonellら[13]は、ある企業から収集された16件のプロジェクトを用いて、前工程の工数から次工程の工数を見積もるモデルの精度を評価している。ただし、開発規模を説明変数として用いておらず、また総工数も見積もっていない。

開発規模を用いた場合とどちらの見積り精度が高いのか、また、開発規模と上流工程工数の両方を用いた場合、更に精度が向上するのか、多重共線性は発生しないのかについては、我々の知る限りこれまで明らかにされていなかった。Yangら[22]は、開発工程の分布は工数見積りに重要であるにもかかわらず、これまで見過ごされがちであったことを指摘している。本研究の主要な貢献は、新たな見積り方法の提案ではなく、上流工程工数を用いると見積り精度を大きく改善できることを示したこと、及び見積りモデル構築における新たなガイドラインを示したことである。ガイドラインは現場への適用が容易な内容であることから、多くのソフトウェア開発企業に適用されることが期待される。

7 おわりに

本研究では、上流工程の活動実績に基づいた工数見積りに着目し、上流工程工数を用いた工数見積りモデルの精度を評価した。上流工程の工数に基づく工数見積りは開発現場で用いられる場合があったが、開発規模を用いたモデルと比較してどちらの見積り精度が高いのかはこれまで明らかではなかった。以下のリサーチ・クエスチョンを設定し、それぞれを分析結果に基づいて明らかにした。

- RQ1: 開発規模のみ、もしくは上流工程工数のみで工数を見積もった場合、どちらの見積り精度が高いのか。

回答:(4.3.1の分析結果より)上流工程工数のみで工数を見積もったほうが見積り精度が高い。

- RQ2: 開発規模、上流工程工数に加え、その他の説明変数を加えた場合、開発規模を含むモデルと上流工程工数を含むモデルでは、どちらの見積り精度が高いのか。

回答:(4.4.1, 4.4.2の分析結果より)その他の説明変数を加えた場合、上流工程工数を含むモデルの見積り精度がわずかに低下する。開発規模とその他の変数を用いたモデルと比較して、計画工数のみを用いたモデルは見積り精度が低く、計画・分析工数のみを用いたモデルは見積り精度が高い。

- RQ3: 開発規模と上流工程工数、両方を用いてモデルを構築した場合、見積り精度はより高まるのか、そのときに多重共線性は発生しないのか。

回答:(4.3.2, 4.4.3の分析結果より)開発規模と上流工程工数、両方を用いてモデルを構築した場合、見積り精度はより高まる。他の説明変数の有無にかかわらず、多重共線性は発生しない。

よって、ソフトウェアの開発規模を正確に計測していない組織では、ある程度の工数見積り精度が期待される、上流工程工数を用いた見積りを適用するとよい。開発規模を正確に計測している組織では、開発規模と上流工程工数の両方を用いることにより、さらなる精度向上が期待される。多重共線性が起きる可能性も低いことから、上流工程工数は工数見積りモデルに積極的に採用すべきであるといえる。

謝辞 本研究の一部は、文部科学省科学研究補助費(基盤C:課題番号25330090,若手A:課題番号24680003)による助成を受けた。

参考文献

- [1] Burgess, C. J. and Lefley, M.: Can genetic programming improve software effort estimation? A

- comparative evaluation, *Information and Software Technology*, Vol. 43, No. 14(2001), pp.863-873.
- [2] Conte, S., Dunsmore, H. and Shen, V.: *Software engineering metrics and models*, Benjamin-Cummings Publishing Co., Inc., Menlo Park, California, 1986.
- [3] International Organization for Standardization and International Electrotechnical Commission: *ISO/IEC 24570:2005 Software engineering - NESMA functional size measurement method version 2.1 - Definitions and counting guidelines for the application of Function Point Analysis*, ISO/IEC, 2005.
- [4] International Software Benchmarking Standards Group (ISBSG): *ISBSG Estimating: Benchmarking and research suite*, ISBSG, 2004.
- [5] International Software Benchmarking Standards Group (ISBSG): *Practical ways to use the ISBSG data*, [http://www.isbsg.org/isbsgnew.nsf/WebPages/EB33DF6DB82ED1C5CA2576DC0081F046/\\$file/Practical%20Ways%20To%20Use%20The%20ISBSG\%20Data.pdf](http://www.isbsg.org/isbsgnew.nsf/WebPages/EB33DF6DB82ED1C5CA2576DC0081F046/$file/Practical%20Ways%20To%20Use%20The%20ISBSG\%20Data.pdf).
- [6] 情報処理推進機構ソフトウェア・エンジニアリング・センター: ソフトウェア開発データ白書 2010-2011, 情報処理推進機構, 東京, 2010.
- [7] 情報処理推進機構: 2012 年度「ソフトウェア産業の実態把握に関する調査」調査報告書, 情報処理推進機構, 東京, 2013.
- [8] 経済調査会経済調査研究所: ソフトウェア開発データリポジトリの分析, 経済調査会, 東京, 2010.
- [9] Kitchenham, B., Pickard, L., MacDonell, S. and Shepperd, M.: What accuracy statistics really measure [software estimation], *Software, IEE Proceedings*, Vol. 148, No. 3(2001), pp. 81-85.
- [10] Kitchenham, B., Pflieger, S., McColl, B. and Eagan, S.: An Empirical Study of Maintenance and Development Estimation Accuracy, *Journal of Systems and Software*, Vol. 64, No. 1(2002), pp. 57-77.
- [11] Lokan, C.: What Should You Optimize When Building an Estimation Model?, in *Proc. of the International Software Metrics Symposium (METRICS)*, 2005, p. 34.
- [12] Lokan, C. and Mendes, E.: Cross-company and single-company effort models using the ISBSG database: a further replicated study, in *Proc. of the International Symposium on Empirical Software Engineering (ISESE)*, 2006, pp. 75-84.
- [13] MacDonell, S. and Shepperd, M.: Using prior-phase effort records for re-estimation during software projects, in *Proc. of the International Software Metrics Symposium (METRICS)*, 2003, pp. 73-86.
- [14] 丸地康平, 今井健男, 岩政幹人, 一條泰男: 工数削減評価のための総工数見積もり方法の提案, ソフトウェア品質シンポジウム (SQiP), 2010.
- [15] Miyazaki, Y., Terakado, M., Ozaki, K. and Nozaki, H.: Robust regression for developing software estimation models, *Journal of Systems and Software*, Vol. 27, No. 1(1994), pp. 3-16.
- [16] Molokken-Østfold, K. and Jørgensen, M.: A comparison of software project overruns - flexible versus sequential development models, *IEEE Transactions on Software Engineering*, Vol. 31, No. 9(2005), pp. 754-766.
- [17] 門田暁人, 小林健一: 線形重回帰モデルを用いたソフトウェア開発工数予測における対数変換の効果, コンピュータソフトウェア, Vol. 27, No. 4(2010), pp. 4.234-4.239.
- [18] 尾崎憲和: 上流工程の問題解決 見積もり編 - 後編, 日経システム構築, 2005, p. 49.
- [19] Shepperd, M. and Schofield, C.: Estimating software project effort using analogies, *IEEE Transactions on Software Engineering*, Vol. 23, No. 12(1997), pp. 736-743.
- [20] 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, 2010.
- [21] Winer, B., Brown, D. and Michels, K.: *Statistical Principles in Experimental Design*, McGraw-Hill, New York, 1991.
- [22] Yang, Y., He, M., Li, M., Wang, Q. and Boehm, B.: Phase distribution of software development effort, in *Proc. of the International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2008, pp. 61-69.



角田 雅照

1997 年和歌山大学経済学部卒業。
2004 年奈良先端科学技術大学院大学
情報科学研究科博士前期課程修了。
2007 年同大学同研究科博士後期課程
修了。同年同大学同研究科特任助教。
2012 年東洋大学総合情報学部助教。
2013 年近畿大学理工学部情報
学科講師。博士(工学)。ソフトウェアメトリクスの研
究に従事。電子情報通信学会, 日本ソフトウェア科学
会, ヒューマンインタフェース学会, 教育システム情
報学会, IEEE 各会員。



戸田 航史

2004 年大阪大学基礎工学部卒業。
2009 年奈良先端科学技術大学院大学
情報科学研究科博士後期課程修了。
2012 年より福岡工業大学情報工学科
助教, 博士(工学)。ソフトウェアメトリクスの研究
に従事。情報処理学会, 電子情報通信学会, IEEE 各
会員。



伏田 享平

2005年大阪府立大学工学部電気電子システム工学科中退。2010年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。同年同大学博士研究員。2011年同特任助教。2012年株式会社NTTデータ入社。現在、技術開発本部プロジェクトマネジメント・イノベーションセンタ主任。博士(工学)。ソフトウェア工学、特にソフトウェア開発データの分析・利用に関する研究に従事。IEEE、情報処理学会 各会員。



亀井 靖高

2005年関西大学総合情報学部卒業。2007年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。2009年同大学院博士後期課程修了。同年日本学術振興会・特別研究員(PD)。2010年カナダ・Queen's大学・博士研究員。2011年九州大学大学院システム情報科学研究院・助教。博士(工学)。ソフトウェアメトリクス、マイニングソフトウェアリポジトリ等の研究に従事。IEEE、IEICE、IPJSJ 各会員。



Meiyappan Nagappan

Meiyappan Nagappan received his PhD from North Carolina State University (NSCU). He is currently a Post Doctoral Fellow at the SAIL lab in Queen's University, Canada. Dr. Nagappan believes that as SE researchers we should look at deriving solutions that encompass the various stakeholders of software systems, and not only software developers. Hence, for the past seven years he has been working on SE research that goes beyond just impacting S/W developers and testers. He has worked on using large scale SE data to also address the concerns of S/W operators, build engineers, and project managers. He continues to publish in top SE venues like TSE, FSE, and has recently won a best paper award at MSR 2012.



鵜林 尚靖

1982年広島大学理学部数学科卒業。1999年東京大学大学院総合文化研究科広域科学専攻広域システム科学系博士課程修了。博士(学術)。1982~2003年(株)東芝に勤務。2003年九州工業大学情報工学部助教授、2010年九州大学大学院システム情報科学研究院教授、現在に至る。2003年度情報処理学会山下記念研究賞受賞。ソフトウェア工学、プログラミング言語モデルに興味を持つ。ACM、IEEE-CS、情報処理学会、電子情報通信学会各会員。