

自主研究

アプリケーション保守費用に 影響を与える要因の分析

アプリケーション保守費用に 影響を与える要因の分析

角田 雅照 奈良先端科学技術大学院大学／近畿大学
松本 健一 奈良先端科学技術大学院大学
大岩 佐和子 押野 智樹 一般財団法人 経済調査会 経済調査研究所 調査研究部 第二調査研究室

1 はじめに

近年、企業活動の推進には情報システムの安定稼働が欠かせないものになっており、アプリケーション保守の重要性が高まっている。一方で情報システムは大規模化、複雑化しており、ユーザだけで保守することができず、ベンダと保守業務の委託契約を結ぶケースも増えている。

保守とは、単にソフトウェア出荷後に発見された欠陥を除去することだけを指すのではない。ソフトウェアは利用されているうちに、ビジネス環境の変化により、機能の拡張、修正が必要となる。保守とは、それらの修正も含んだ概念であり、JIS X 0161 [1] では以下の4つに分類している。

- (1) 是正保守：ソフトウェア製品の引渡し後に発見された問題を訂正するためにおこなう受身の修正
- (2) 予防保守：引渡し後のソフトウェア製品の潜在的な障害が顕在化する前に発見し、是正を行うための修正
- (3) 適応保守：引渡し後、変化した又は変化している環境において、ソフトウェア製品を使用できるように保ち続けるために実施するソフトウェア製品の修正
- (4) 完全化保守：引渡し後のソフトウェア製品の性能又は保守性を改善するための修正

これまでも保守業務の現状分析など、保守に関する研究を行ってきたが [2] [3]、本稿はそれを更に進めたものである。

本稿の目的は、保守費用の妥当性を検討する際に参考となる資料を提供することである。そのために、保守の作業効率に影響する要因と、保守の技術者単価に影響する要因を分析した。作業効率は、作業時間あたりのソフトウェアの修正量（プログラム本数）を表す。

技術者単価は作業時間あたりの保守技術者の費用を指す。以降の分析において、要因別の作業効率と技術者単価を示す。年間のソフトウェア保守量がおおむね決まっているならば、これらを用いて年間の作業時間と費用が推定できる。仮に「金融業のソフトウェアでは、作業効率が0.05本/時（要因別の作業効率）」とすると、年間の保守プログラム本数が10本ならば、作業時間は200時間と推定できる。さらに、仮に「金融業での技術者単価（時間単価）が10,000円（要因別の技術者単価）」とすると、保守費用は、作業時間と技術者単価の積より200万円と推定できる。

2 分析対象のデータ及び分析方法

本稿では、一般財団法人経済調査会によって2006年から2016年にかけて収集された836件のアプリケーション保守の事例を分析対象とした。これらの事例から、データの信頼性や分析結果の一貫性を考慮し、以下の条件で107件のデータを抽出した。

- 年間保守契約費用が記録されており、その金額が100万円以上である。
- 年間保守作業時間（延べ）が記録されている。
- 是正保守、予防保守などの各保守の作業比率が記録されている。

作業効率の分析にあたっては、予備分析の結果から、プログラム修正効率を指標とした。プログラム修正効率は、以下の式により算出する。

プログラム修正効率 = 修正プログラム本数 ÷ 作業時間

作業効率や技術者単価に影響を与える要因として、システム構成や業種などの事例ごとの属性を分析対象とした。作業効率・技術者単価と各要因との関連の強さの分析で、主に相関比を用いた。これは、業種など

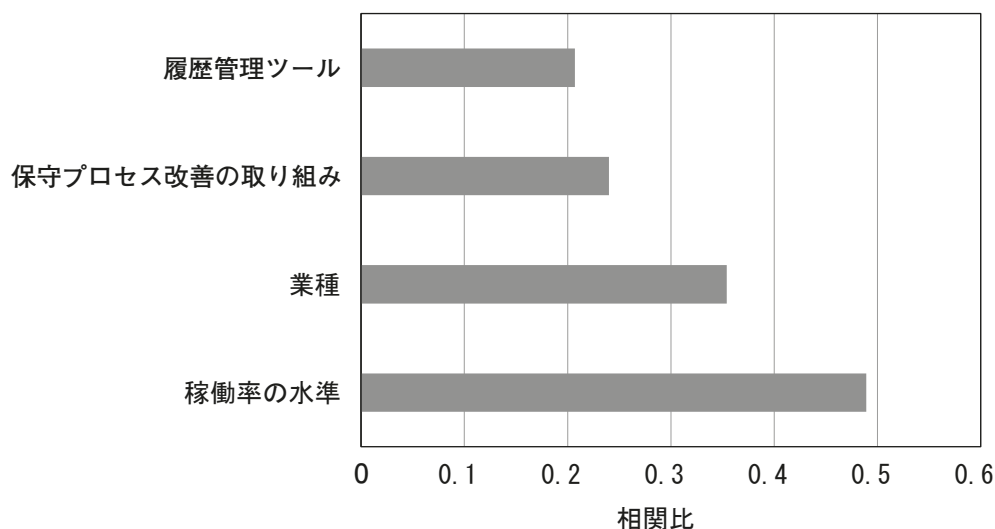
のカテゴリで表される要因と作業効率などの数値との関係の強さを表す指標である。0から1の値を取り、値が大きいほど関連が強いことを示す。本稿では相関比が0.2を上回っている関係について詳細に分析した。また、数値で表される項目同士、例えば作業効率と技術者単価の関連の分析では、順位相関係数を用いて分析した。これは-1から1の値を取り、絶対値が大きいほど関連が強いことを示す。こちらについても絶対値が0.2を上回っている関係について着目した。

その他に、データの分布の分析には箱ひげ図を用いた。箱ひげ図とは、四分位数を用いてデータの分布を表す図である。箱の下辺は、例えば100個の値を昇順に並べた場合に25番目に現れる値(第1四分位)を示

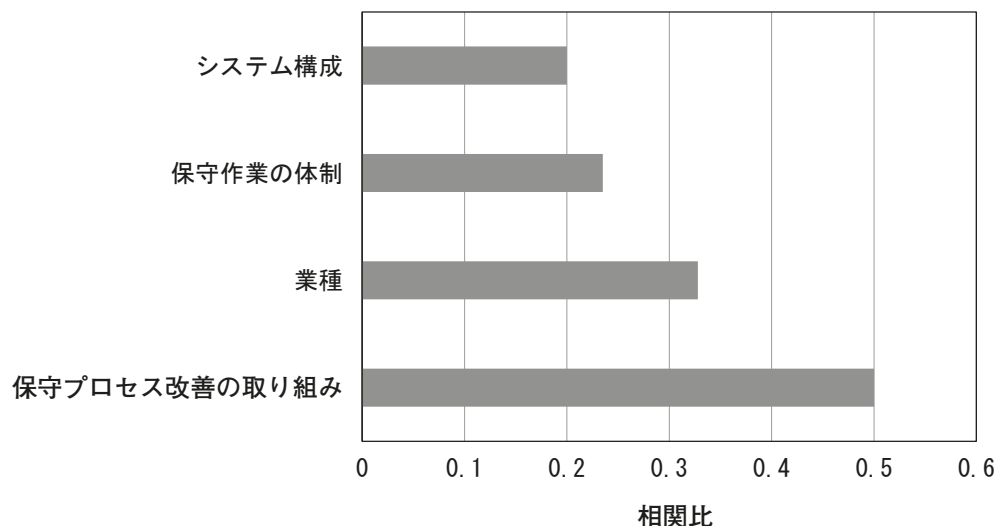
し、上辺は75番目に現れる値(第3四分位)を示す。箱の中の太線は中央値(第2四分位)を表す。図中のひげの部分は、それぞれ箱の長さの1.5倍を超えない範囲にある最小値、最大値を示し、丸印は箱の上辺下辺から箱の長さの1.5倍以上離れた値、星印は箱の上辺下辺から箱の長さの3倍以上離れた値を示す。箱の部分に全体の50%のデータが含まれる。図を見やすくするために、値が極端に大きいデータの一部を除外した。

作業効率と各要因との関連の強さを図表1に、技術者単価と各要因との関連の強さを図表2に示す。以降では、これらの図表で比較的作業効率や技術者単価との関連が明確であったものを中心に、詳細に分析した結果を述べる。

図表1 作業効率と各要因との関連の強さ



図表2 技術者単価と各要因との関連の強さ



3 作業効率への影響要因

3.1 保守プロセス改善の取り組み

企業によっては、保守の作業効率の改善や品質向上のために、専任の担当者を置いてプロセス（工程、作業手順）の改善に取り組んでいる場合がある。この取り組みの状況が保守作業の効率に影響している可能性がある。

分析対象データには、保守改善のための担当者に関して、次のいずれかの回答が記録されている。

1. 改善専任の担当がいる
2. 専任ではないが改善担当がいる
3. プロセス改善の担当はいない
4. その他

この属性に基づいて分析したプログラム修正効率の箱ひげ図を図表3に示す。「改善専任の担当がいる」の場合、「専任ではないが改善担当がいる（改善担当者）」の場合よりも箱の位置（データの上位25%から下位25%までの分布）が高めとなっている。これは、「改善専任の担当がいる」の上位25%（箱の上部）の効率が高く、かつ、下位25%（箱の下部）の効率も高いことを示している。「プロセス改善の担当はいない」の場合、箱が他より大きい。これは、データにばらつきがあることを示し、プロセス改善の担当がない場合、他の場合と比較して事例ごとの作業効率の違いが大きいことを示している。

- プロセス改善の担当者の有無は、作業効率に関連している可能性がある。
- プロセス改善の担当者がいない場合、事例によって作業効率が大きく異なる。

3.2 業種

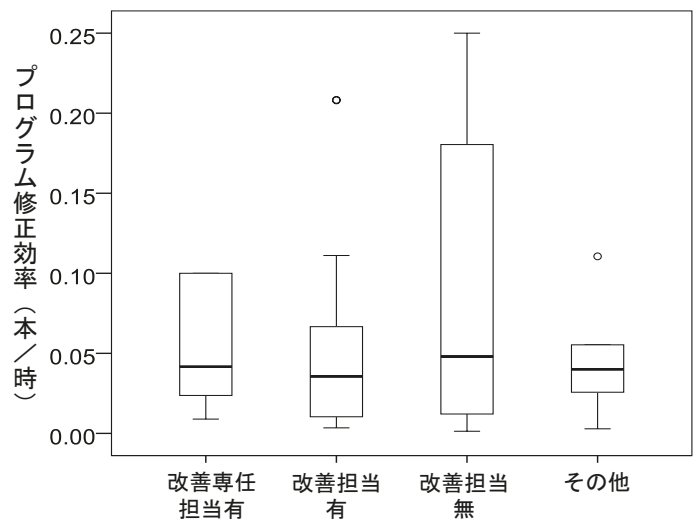
分析対象のデータは多様な組織から収集されており業種も多様である。業種が異なれば、作業効率が異なる可能性がある。例えば、信頼性の求められる金融業の情報システムの場合、欠陥の発生を抑えるために、テストを厳重に行う必要がある。その場合、作業時間

あたりのソフトウェア修正量は他の業種より小さくなり、作業効率が低くなる可能性がある。

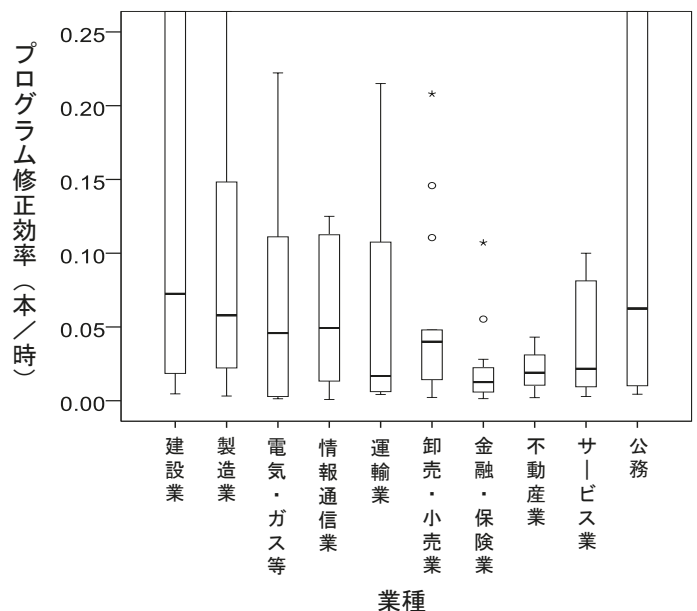
業種別のプログラム修正効率の箱ひげ図を図表4に示す。ここでは、データの件数が8件以上の業種を取り上げる。「建設業」「製造業」の効率が低い傾向があり、「金融・保険業」の効率が低い傾向があった。

- 建設業、製造業は保守の作業効率が高い傾向があり、金融・保険業は低い傾向が見られる。

図表3 プロセス改善の取り組みとプログラム修正効率の関係



図表4 業種とプログラム修正効率の関係



3.3 稼働率の水準

稼働率はソフトウェアが稼働している時間の割合であり、以下の式で定義される。

$$\text{稼働率} = (\text{運用時間} - \text{障害停止時間}) \div \text{運用時間}$$

ソフトウェアの停止が業務に与える影響が大きいほど、求められる稼働率の水準が高くなる。稼働率の要求水準が高いソフトウェアの場合、保守時に欠陥の発生を抑えるために、テストに、より多くの時間をかけ、その結果、作業効率が低下する可能性がある。

分析対象データには、要求される稼働率の水準として、次のいずれかの回答が記録されている。

1. 95%以下
2. 95%を超え99%以下
3. 99%を超え99.9%以下
4. 99.9%を超え99.99%以下
5. 99.99%を超え99.999%以下
6. 99.999%超
7. その他

稼働率の水準とプログラム修正効率の箱ひげ図を**図表5**に示す。稼働率の水準が「99%を超え99.9%以下」や「99.9%を超え99.99%以下」の場合と比較して、「99.999%超」の場合は箱の位置が低くなっている。

- 稼働率の要求水準が高い場合、作業効率が低くなる傾向が見られる。

3.4 履歴管理ツール

ソフトウェアの保守では、履歴管理ツールを用いている場合がある。履歴管理ツールとは、保守開発の履歴の検索を支援するツールを指す。過去の保守対応で、今回の保守対応と類似のものがあり、その時の資料を参照することができれば、作業効率を改善できたり欠陥発生を抑えられたりする可能性がある。

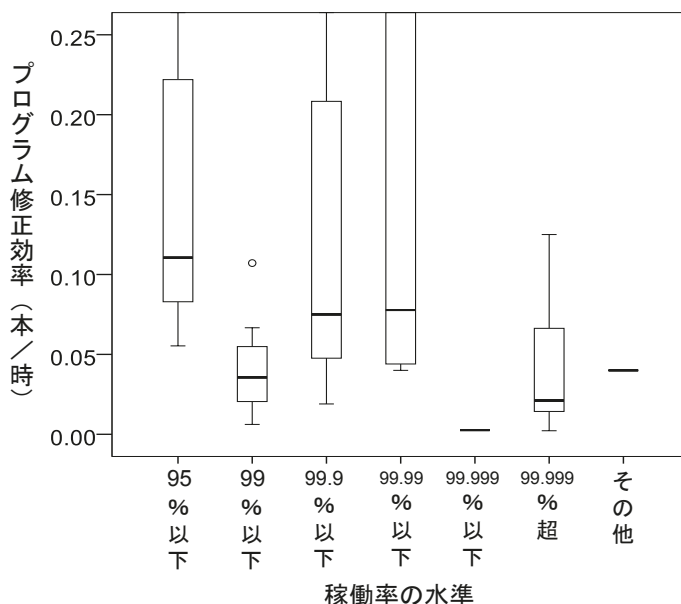
履歴管理ツールの使用の有無とプログラム修正効率の関係を示す箱ひげ図を**図表6**に示す。

効率の中央値や下位25%については、履歴管理ツール導入の有無による差はなかった。

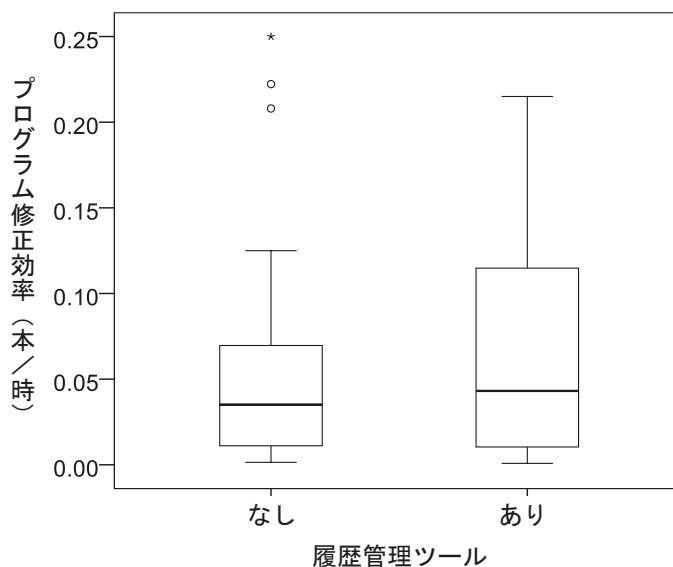
ただし、効率の上位25%については、履歴管理ツールを導入しているほうの値が高かった。これは、ツールを導入することにより、必ずしも効率が向上するわけではないが、効率が高い事例もある程度見られることを示している。

- 履歴管理ツールを導入している場合、必ずしも作業効率が高まるわけではないが、事例によっては高まる傾向が見られる。

図表5 稼働率の水準とプログラム修正効率の関係



図表6 履歴管理ツールとプログラム修正効率の関係



3.5 品質との関連

2章の作業効率と各要因との関連の強さの分析では抽出されなかったが、作業効率には、品質（ソフトウェアの欠陥数や不具合数）の要求度合いが影響している可能性がある。そこで品質と作業効率の関連を分析した。

まず、年間の不具合発生件数と作業効率との順位相関係数を算出した。不具合発生件数は不具合の重大度ごとに記録されている。結果を図表7に示す。

順位相関係数が正の値で大きなもの（0.3を超えるもの）はなかった。このため、作業効率が高いほど不具合発生件数も増加するという傾向はないといえる。従って、「作業効率が高い事例は、テストなどの時間を短縮して効率を高めた結果である」という可能性は低いといえる。

- 作業効率と保守ソフトウェアの品質との関連性は低い。

図表7 作業効率と年間不具合発生件数との順位相関係数

	プログラム修正効率
重大不具合発生件数	-0.10
中程度不具合発生件数	-0.02
軽微不具合発生件数	0.10

4 技術者単価の影響要因

4.1 保守プロセス改善の取り組み

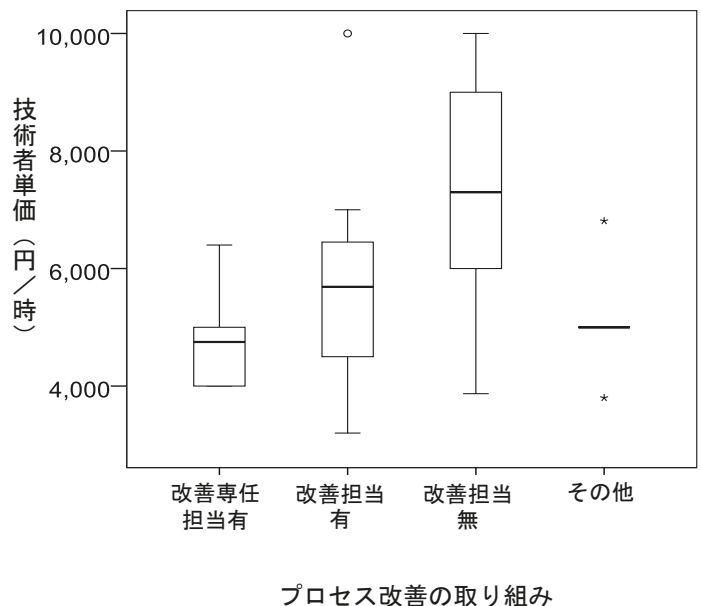
3.1節で説明した保守改善のための担当者に関する回答と、技術者単価との関連を分析した。保守改善のための担当者の有無の違いが、技術者単価に直接影響することは考えにくい。これは、改善担当者が保守作業に直接関与することはないと思われるためである。ただし、保守プロセス改善に積極的に取り組んでいる（改善の担当者がある）場合、プロセス改善が進んでおり、その結果、作業手順が整理され比較的スキルの高くない技術者による保守が可能となっている可能性もある。そこで、保守改善のための担当者に関する属性（3.1節参照）と技術者単価との関係を分析した。

プロセス改善の取り組みの状況と技術者単価との関

係を示す箱ひげ図を図表8に示す。「改善専任の担当がいる」の回答は6件と少なかったため、「専任ではないが改善担当がいる（改善担当有）」と「プロセス改善の担当はいない」に着目する。前者のほうが技術者単価の中央値が小さく、かつ箱の位置も前者のほうが低かった。

- プロセス改善の担当者がある場合、技術者単価が低い傾向が見られる。

図表8 プロセス改善の取り組みと技術者単価の関係



4.2 業種

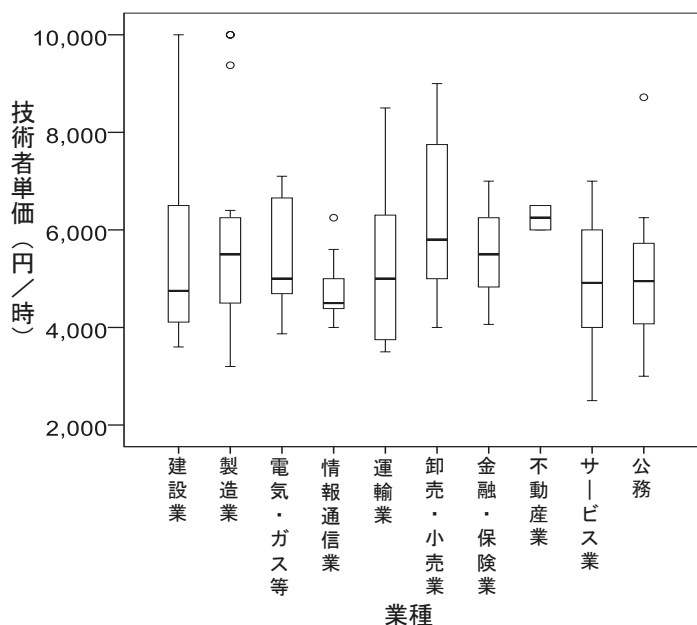
3.2節で述べたように、保守対象のソフトウェアが稼働している組織の業種は多様である。業種によって、保守技術者に求められる業務知識や技術が異なる可能性があり、それらが技術者単価に影響していることが考えられる。

業種別の技術者単価の箱ひげ図を図表9に示す。

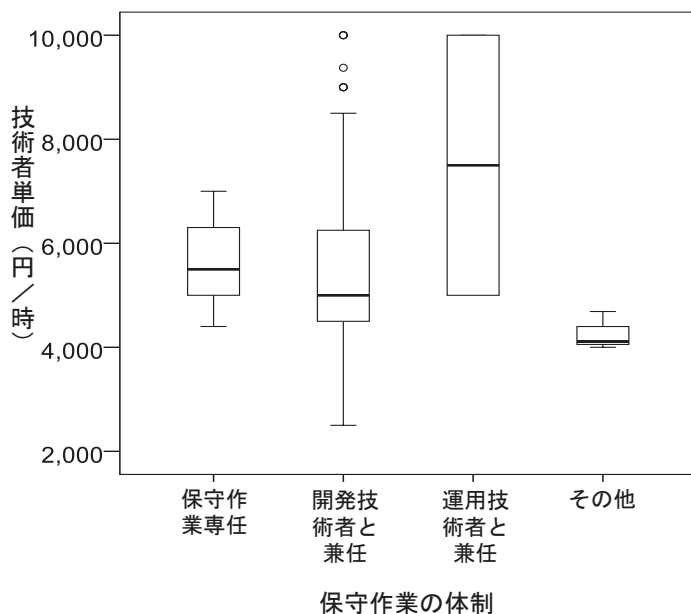
ここでは、箱ひげ図において単価が低めまたは高めの傾向が見られ、かつデータ件数が少なくない業種に着目する。「建設業」、「情報通信業」については、単価の中央値が他の業種と比べて低い傾向が見られた。一方、単価が高めの傾向が見られるという業種はなかった。

- 建設業、情報通信業については、技術者単価が比較的低い傾向が見られる。

図表9 業種と技術者単価の関係



図表10 保守作業の体制と技術者単価の関係



4.3 保守作業の体制

ソフトウェアの保守では、技術者が保守専任の場合とそうでない場合、例えば、ソフトウェア開発技術者と兼任の場合がある。そこで保守作業の体制と技術者単価との関係を分析した。

分析対象データには、保守作業の体制について、次のいずれかの回答が記録されている。

1. 保守技術者は保守作業専任
2. 保守技術者は開発技術者と兼任
3. 保守技術者は運用技術者と兼任
4. その他

保守作業の体制別の技術者単価の箱ひげ図を図表10に示す。

「保守技術者は運用技術者と兼任」と「その他」の回答は少なかったため、これらは考慮せずに分析する。「保守作業専任」のほうが「開発技術者と兼任」よりも単価の中央値が大きく、かつ下位25%の単価も高かった。

- 保守技術者が保守作業専任の場合、開発技術者と兼任の場合と比較して、技術者単価が高い傾向がある。

4.4 システム構成

保守対象のソフトウェアが稼働するシステムの構成として、クライアントサーバやメインフレームなどがある。システム構成が異なると、OS(オペレーティングシステム)やプログラミング言語、利用するツールなどが異なるため、保守技術者に求められるスキルも異なる。保守技術者の持つスキルによって技術者単価も異なると考えられるため、システム構成は単価に影響する可能性がある。そこでシステム構成と技術者単価との関連を分析した。

分析対象データには、システム構成について、次のいずれかの回答が記録されている。なお、ソフトウェアが複数のシステムで用いられている場合、回答として複数のシステム構成が選択されることがある。

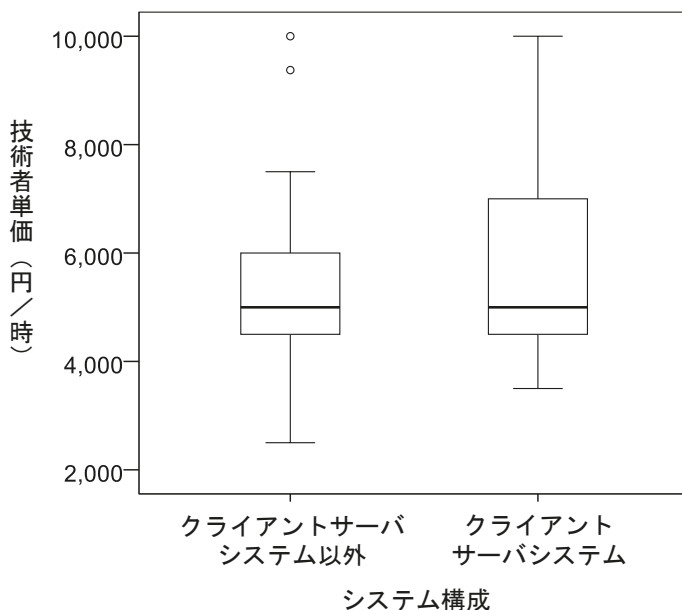
1. クライアントサーバシステム
2. Web系システム
3. メインフレームシステム
4. その他

ここでは技術者単価と関連が見られたクライアントサーバシステムのみ取り上げる。

システム構成がクライアントサーバであるか否かによる技術者単価の箱ひげ図を**図表11**に示す。単価の中央値に差はないが、クライアントサーバはデータのばらつきが大きく、かつ上位25%の単価が高くなっていた。

- クライアントサーバシステムの場合、技術者単価が高くなる傾向がある。

図表11 システム構成と技術者単価の関係



4.5 保守作業比率との関連

1章で述べたように、保守には4つの種類がある。それぞれの実施方法は異なり、技術者に求められるスキルも異なる。例えば、是正保守はプログラムを十分理解できる能力が必要であり、経験豊富な単価の高い技術者が求められる可能性がある。そこで下記の4つの種類ごとの作業比率と、技術者単価との関連を分析した。作業比率はそれぞれの種類の保守が占める割合を示し、これらを合計すると100%となる。

- 是正保守比率
- 予防保守比率
- 適応保守比率
- 完全化保守比率

これらの作業比率と技術者単価との順位相関係数を**図表12**に示す。是正保守と適応保守については、他に比べ相関が大きかった。

図表12 保守作業比率と技術者単価の順位相関係数

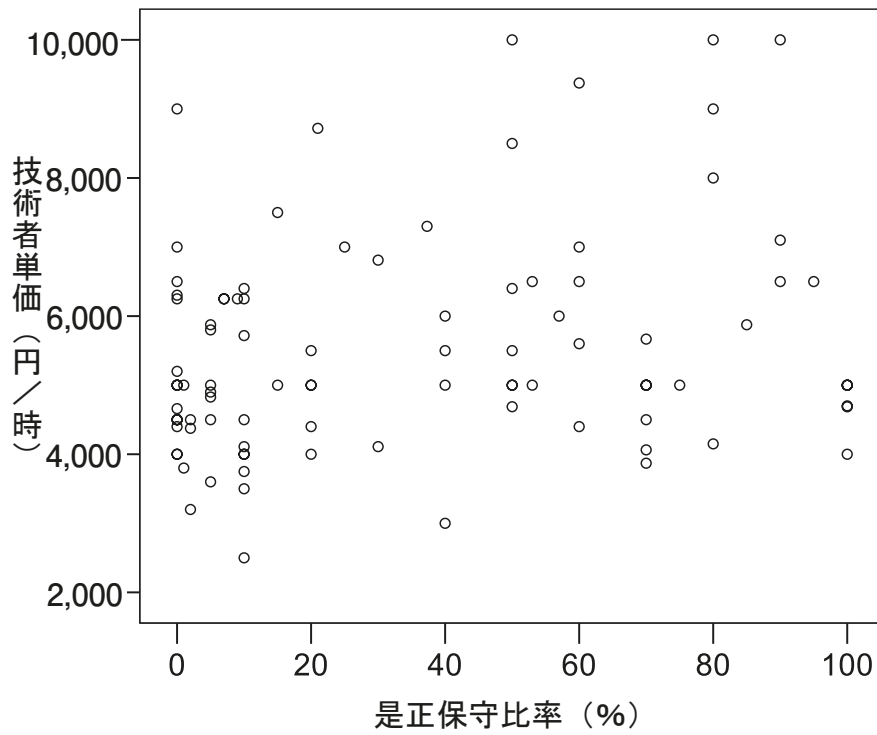
	技術者単価
是正保守比率	0.220
予防保守比率	0.060
適応保守比率	-0.278
完全化保守比率	0.033

是正保守比率と技術者単価の関係を示す散布図を**図表13**に示す。是正保守がない(0%)場合(**図表13**の左側)と比べて、是正保守の割合が比較的高い(40%以上)場合のほうが、若干技術者単価が高い傾向が見られた。

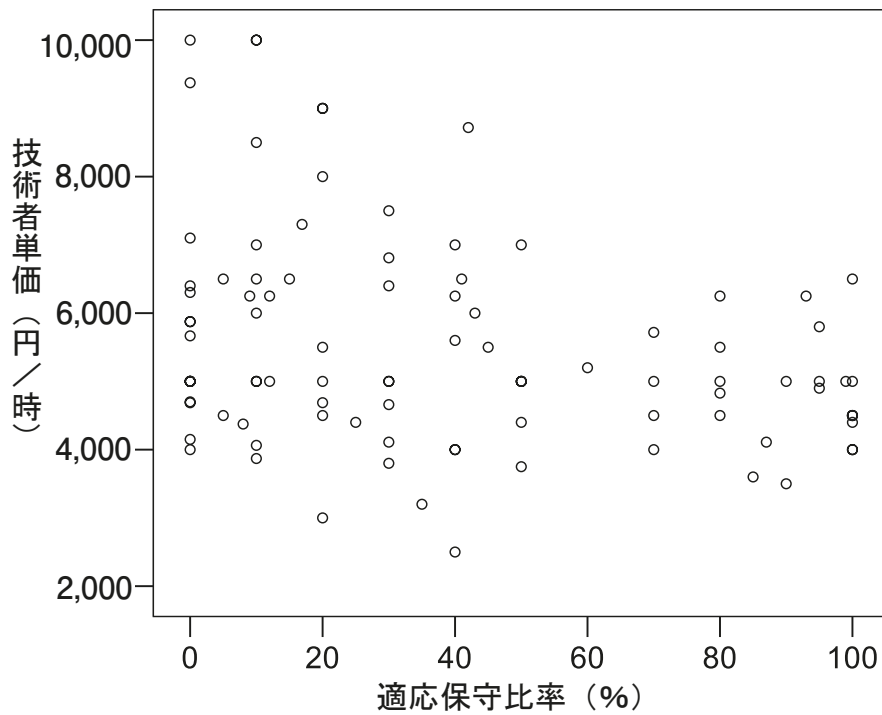
適応保守比率と技術者単価との散布図を**図表14**に示す。適応保守の割合が比較的低い(40%以下)場合、技術者単価が高い事例がいくつか見られた。適応保守の比率が高いほど技術者単価が低い傾向が見られた理由は、他の保守と比較して特別な技術を要しないためである、または、特定の業種で適応保守が多いなど、が考えられる。

- 是正保守の比率が高い場合、技術者単価が高い傾向が若干見られる。
- 適応保守の比率が高い場合、技術者単価が低い傾向が若干見られる。

図表13 是正保守比率と技術者単価と関係



図表14 適応保守比率と技術者単価の関係



4.6 品質との関連

品質と作業効率については3.5節で分析したが、品質と技術者単価にも関係がある可能性がある。例えば、技術者単価が低い場合、技術が低いために欠陥を発生させやすい可能性もある。そこで品質と技術者単価との関連を分析した。

年間の不具合発生件数と技術者単価との順位相関係数を図表15に示す。順位相関係数が負の値で絶対値が大きなもの(-0.1を下回るもの)はなかった。

このため、技術者単価が低いほど不具合件数も増加するという傾向はないといえる。

- 技術者単価とソフトウェアの品質との関連性は低い。

図表15 技術者単価と年間不具合発生件数との順位相関係数

	技術者単価
重大不具合発生件数	0.16
中程度不具合発生件数	-0.01
軽微不具合発生件数	0.01

5 保守作業時間・保守費用の予測

本章では年間保守作業時間と年間保守費用の予測を試みた。なお、本分析では予測がどの程度の精度で可能であるかは示すが、結果の濫用を防ぐために、実際の予測モデルは提示しない。

5.1 保守作業時間の予測

修正プログラム本数などから、保守に必要な作業時間が予測できれば、保守費用の妥当性の検討に利用できる。例えば、業種が製造業であり、年間で修正する予定のプログラムがおおむね10本程度であるとする。このとき、製造業のソフトウェアで10本プログラムを修正するには作業時間が100時間かかると予測され

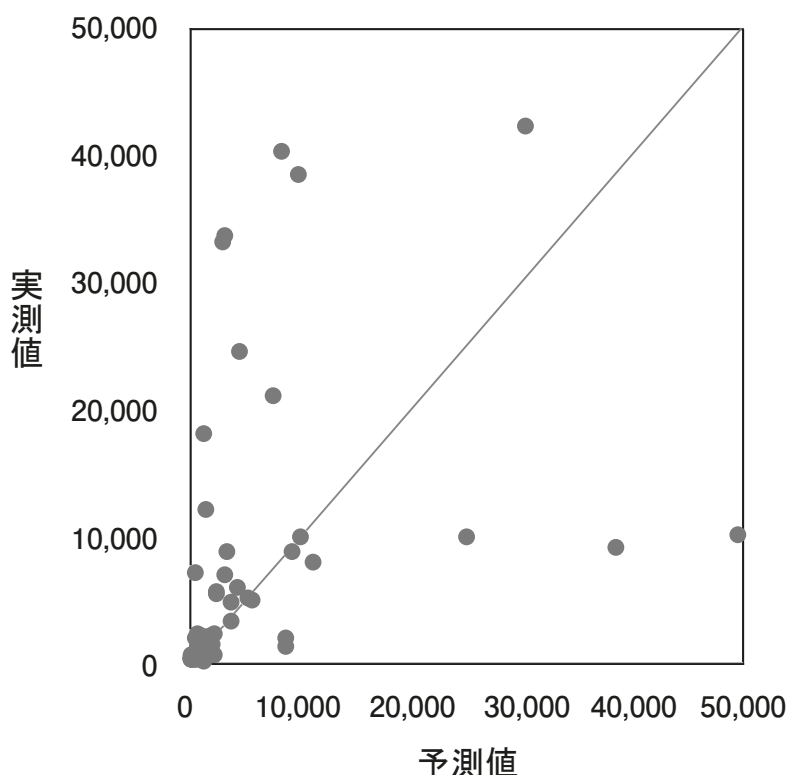
れば、100時間に技術者単価を乗じた金額が、保守費用として妥当であると判断できる。

この例での100時間という予測作業時間は、過去の事例に基づき、統計的に算出されるものであるが、天気予報などと同様に誤りが含まれる。例えば、過去の事例10件のうち、業種が製造業、修正プログラムが10本の場合、100時間の事例が5件、150時間の事例が3件、50時間の事例が2件だとすると、100時間という予測結果は必ずしも正しくない。

このような、予測に含まれる誤りは「相対誤差」により表現される。 $|\text{予測値} - \text{実測値}| \div \text{予測値}$ で計算され、例えば、予測値が100時間、実際の値(実測値)が150時間とすると $|\text{予測値}100\text{時間} - \text{実測値}150\text{時間}| \div \text{予測値}100\text{時間} = 50\%$ となる。相対誤差が小さいほど、予測結果が信頼できるといえる。

本節では、重回帰分析という手法を用いて、統計的に作業時間を予測し、誤差の算出には交差検証と呼ばれる方法を用いた。ここでの作業時間は委託側と受託側の作業時間の合計を指し、モデルの構築には予測に必要なデータが記録されている65件の事例を用いた。相対誤差の中央値は111%、相対誤差の平均値は325%となった(R^2 は0.72、調整済 R^2 は0.69)。誤差の程度を表すために、予測値と実測値の散布図を図表16に示す。図の対角線上に点が集まっているほど、予測結果が正しいことを示す。図では対角線上にない点も多く、予測精度は十分高いとはいえない。このときの説明変数を図表17に示す。説明変数とは、作業時間を予測するために必要なデータを指し、表の「標準化偏回帰係数」が大きいほど、作業時間(作業効率)への影響が大きいことを示す。値が正の場合、作業時間を増やす効果があり、負の場合は時間を減らす効果がある。変数選択という方法により、作業時間への影響が小さい項目は除外されている。当然ではあるが、図表17よりプログラム修正本数が作業時間に最も影響が大きいことがわかる。

図表16 年間保守作業時間の予測誤差



図表17 年間保守作業時間の予測モデル

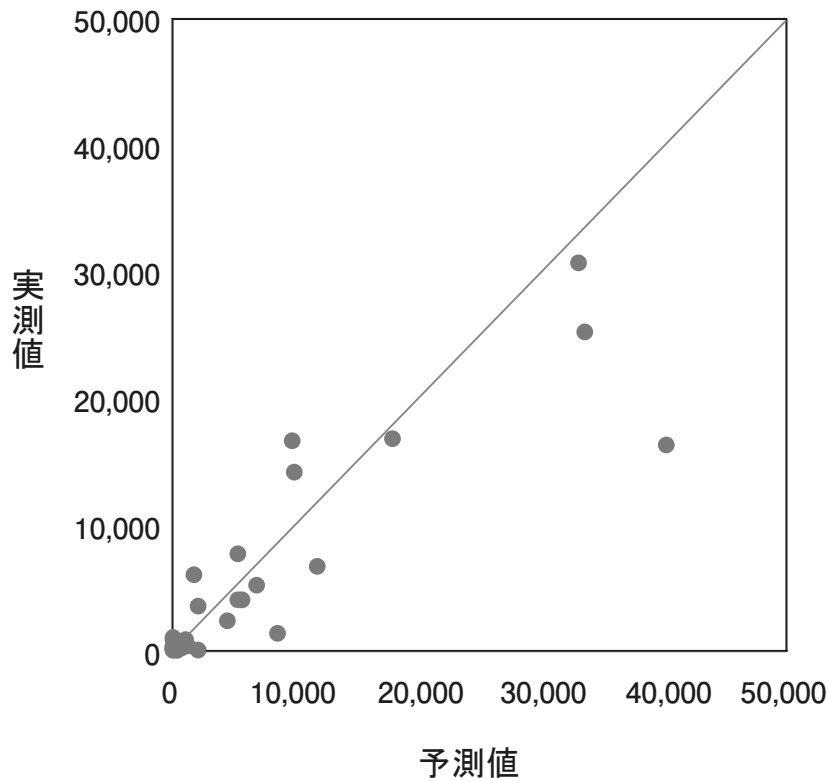
説明変数	標準化偏回帰係数
プログラム修正本数	0.43
バッチプログラム修正本数	0.29
構成管理ツール	0.26
完全化保守比率	0.22
保守用ドキュメント	0.12
予防保守比率	-0.11

次に、「保守プロセス改善の取り組み」の状況が記録されている事例のみを用いて予測を行った。この場合、予測対象の事例は35件、相対誤差中央値は109%、相対誤差平均値は214%となり、誤差が改善した (R^2 は0.78、調整済み R^2 は0.74)。予測値と実測値の散布図を図表18に示す。散布図の対角線上に点が比較的集まっており、誤差が小さくなっていることがわかる。この場合の標準化偏回帰係数を図表19に示す。

保守プロセス改善の取り組みの状況について回答できる組織については、収集されたデータの正確性が高く、そのために予測時の誤差が小さくなった可能性がある。

- 保守プロセス改善の取り組みが記録されている事例では、かなり大まかにではあるが作業時間を予測することができる。

図表18 年間保守作業時間の予測誤差(プロセス改善の取り組み状況の記録あり)



図表19 年間保守作業時間の予測モデル(プロセス改善の取り組み状況の記録あり)

説明変数	標準化偏回帰係数
プログラム修正本数	0.47
予防保守比率	0.37
バッチプログラム修正本数	0.39
完全化保守比率	0.20
プロセス改善取り組み	0.17

5.2 保守費用の予測

前節では年間保守作業時間の予測を行ったが、本節では年間保守費用そのものを直接予測することを行った結果について述べる。前節と同様に重回帰分析を用いて、年間保守費用を予測した。モデルの構築には予測に必要なデータが記録されている36件の事例を用いた。

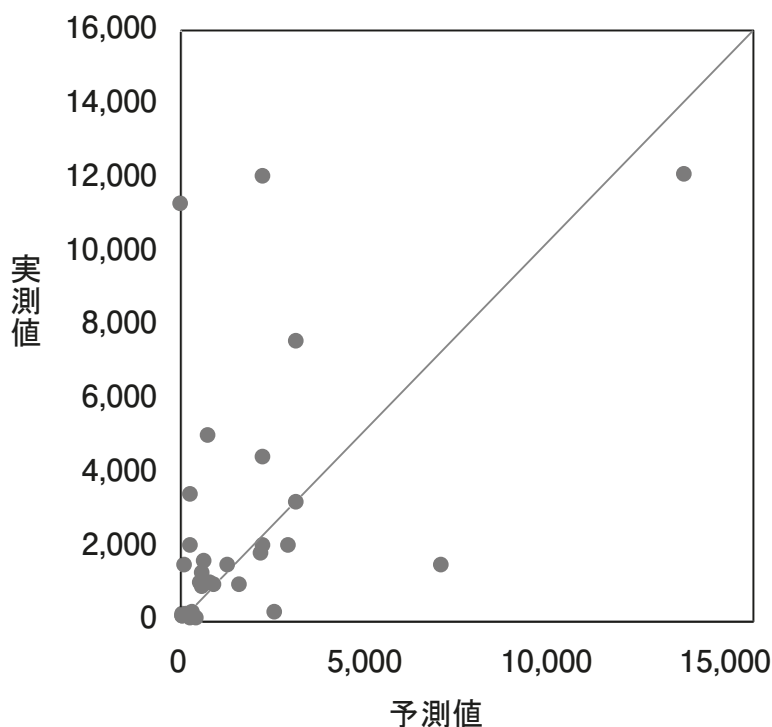
相対誤差の中央値は129%、相対誤差の平均値は18044%となった(R^2 は0.64、調整済 R^2 は0.58)。誤差の大きいデータ1件を取り除くと、相対誤差の中央値は125%、相対誤差の平均値は304%となった。予測値と実測値の散布図を図表20に示す。予測値と実測値との差は極端に大きくはなかったため、ある程度

の精度で保守費用を予測できているといえる。

説明変数と標準化偏回帰係数を図表21に示す。変数選択の結果、「委託者と受託者の作業時間比率」が説明変数に残らなかった。通常は、この作業比率が変われば保守費用は変わると考えられる。例えばプログラム10本を修正する場合、受託者の作業比率が100%の場合と50%の場合とでは、後者のほうが受託者の作業時間が短くなり、保守費用も小さくなるのが通常である。このため、予測誤差は大きくないが、この結果の利用には十分な検討が必要であるといえる。

- 年間保守費用は、大まかに予測できるが、実際に利用するにはさらなる検討が必要である。

図表20 年間保守費用の予測誤差



図表21 年間保守費用の予測モデル

説明変数	標準化偏回帰係数
プログラム修正本数	0.29
保守プロセスの標準化状況	-0.32
完全化保守比率	0.32
Web系システム	0.34
バッチプログラム本数	0.26

6 おわりに

最後に、ソフトウェア保守費用の妥当性の参考資料として、どのように分析結果を利用すればよいかを述べる。まず、作業時間の推定手順は以下ようになる。

1. 2章で示した要因のうち、寄与率が高い項目（業種など）に着目する。
2. 寄与率が高い項目について、自組織に当てはまっているカテゴリ（例えば業種なら金融業など）に着目する。
3. 手順2のカテゴリの箱ひげ図の箱が大きい場合、手順1に戻る。
4. 箱ひげ図から作業効率の中央値を読み取る。
5. 年間のプログラム修正本数を調べ、手順4で確認した作業効率で除し、作業時間を推定する。

次に、技術者単価と保守費用の推定手順は以下のようになる。手順1から3は作業時間の推定手順と同様である。

4. 箱ひげ図から技術者単価の中央値を読み取る。
5. 1から3の手順で推定した作業時間から、委託者側の作業時間を減じる。
6. 手順5で求めた作業時間に、手順4の技術者単価を乗じる。

なお、各箱ひげ図からもわかるように、同じ条件（例えば業種が金融業）でも事例によって作業効率や技術者単価が異なるため、推定結果を絶対視するのではなく、参考として利用することが望ましい。

参考文献

- [1] 日本規格協会：JIS X0161 ソフトウェア技術—ソフトウェアライフサイクルプロセス—保守、日本規格協会、2008.
- [2] 経済調査研究レビューVol.13「ソフトウェア保守改善に向けたデータ分析」、角田雅照、門田暁人、松本健一、大岩佐和子、押野智樹、pp.100-110、2013.
- [3] 経済調査研究レビューVol.03「ソフトウェア業務の現状分析」、角田雅照、門田暁人、松本健一、pp.46-59、2008.