

February 2007

Systematische Auswahl einer Methode zur Aufwandschätzung in der Softwareentwicklung

Ulrike Dowie

Universität Stuttgart, dowie@wi.uni-stuttgart.de

Georg Herzwurm

Universität Stuttgart, herzwurm@wi.uni-stuttgart.de

Follow this and additional works at: <http://aisel.aisnet.org/wi2007>

Recommended Citation

Dowie, Ulrike and Herzwurm, Georg, "Systematische Auswahl einer Methode zur Aufwandschätzung in der Softwareentwicklung" (2007). *Wirtschaftsinformatik Proceedings 2007*. 73.

<http://aisel.aisnet.org/wi2007/73>

This material is brought to you by the Wirtschaftsinformatik at AIS Electronic Library (AISeL). It has been accepted for inclusion in Wirtschaftsinformatik Proceedings 2007 by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

In: Oberweis, Andreas, u.a. (Hg.) 2007. *eOrganisation: Service-, Prozess-, Market-Engineering*; 8. Internationale Tagung Wirtschaftsinformatik 2007. Karlsruhe: Universitätsverlag Karlsruhe

ISBN: 978-3-86644-094-4 (Band 1)

ISBN: 978-3-86644-095-1 (Band 2)

ISBN: 978-3-86644-093-7 (set)

© Universitätsverlag Karlsruhe 2007

Systematische Auswahl einer Methode zur Aufwandschätzung in der Softwareentwicklung

Ulrike Dowie, Georg Herzwurm

Lehrstuhl für Wirtschaftsinformatik (Unternehmenssoftware)

Universität Stuttgart

70174 Stuttgart

{dowie,herzwurm}@wi.uni-stuttgart.de

Abstract

Während sowohl in wissenschaftlichen Arbeiten als auch in der Praxis des Projektmanagements Einigkeit über die Wichtigkeit der Aufwandschätzung herrscht, ist die Diskrepanz bezüglich der Verbreitung verschiedener Methoden zwischen Wissenschaft und Praxis erheblich.

Dieser Beitrag schildert einen systematischen Ansatz zur Auswahl geeigneter Aufwandschätzmethoden in der Softwareentwicklung. Hierbei liefern wissenschaftliche Arbeiten Informationen über verfügbare Methoden sowie deren Stärken und Schwächen, während aus Sicht der Praxis Anforderungen formuliert werden, die an jegliche Aufwandschätzmethode gestellt werden.

Die Anwendbarkeit des Ansatzes wird anhand eines Beispiels erläutert. Gleichzeitig wird gezeigt, wie verschiedene Schätzmethode integriert werden können.

1 Einleitung

In der Vielzahl wissenschaftlicher Veröffentlichungen, die sich mit Aufwandschätzmethoden befassen, gibt es kaum Anhaltspunkte, welche Methode in einer gegebenen Situation zu wählen ist [MySS05, 380]. Im Folgenden wird ein Ansatz erläutert, aus der Menge möglicher Schätzmethoden eine Auswahl zu treffen, die sowohl Bedürfnisse als auch Fähigkeiten einer softwareentwickelnden Organisation berücksichtigt [MaSh03, 1].

1.1 Bedeutung der Auswahl von Aufwandschätzmethoden

Aufwandschätzung ist als wichtige Aufgabe des Projektmanagers als Bestandteil der Projektsteuerung anerkannt [BuFa04]. Sie dient bspw. zur Ressourcenallokation, Zeitplanung, Angebotserstellung bei Auftragsentwicklung, sowie als Entscheidungsgrundlage beim Abwägen zwischen Qualitätszielen, Time to Market und Produktumfang [BuFa04].

Hierzu existiert eine Vielzahl von Methoden [vgl. zum Überblick Kitc96; MKLP00]. Häufige Widersprüche zwischen Ergebnissen verschiedener Untersuchungen zeigen, dass keine Methode als generell gut oder schlecht beurteilt werden kann [MaSh03; MySS05]. Die Eignung einer Schätzmethode ist offenbar von situativen Faktoren wie den Zielen und Rahmenbedingungen abhängig, die in einer konkreten Schätzsituation vorliegen.

In der Softwareentwicklungspraxis findet jedoch kaum eine Analyse dieser Faktoren und dementsprechende Auswahl geeigneter Schätzmethoden statt. Meist wird verwendet, wozu in der Organisation Wissen und Erfahrung verfügbar sind [Molø04].

Wissenschaftliche Arbeiten sind meist der Entwicklung und Untersuchung jeweils einer Methode gewidmet (z. B. [Lang04]) oder vergleichen einige ausgewählte Methoden hinsichtlich der Schätzgenauigkeit anhand einer bestimmten Datenbasis (von Autoren ermittelte [WaJe98] oder öffentlich verbreitete empirische Datenbestände [MKLP00; KCCS00]).

Sollen zusätzliche Anforderungen an die Methode, weitere Methoden oder andere bzw. eigene historische Daten berücksichtigt werden, sind derartige wissenschaftliche Erkenntnisse wenig hilfreich. Daher ist sowohl aus Sicht der Praxis als auch der Wissenschaft zu untersuchen, welche Schätzmethode in einer bestimmten Schätzsituation a priori auszuwählen ist [MaSh03].

1.2 Ziel und Vorgehensweise

Die Ziele dieses Beitrags sind folgende:

- Ein Konzept zur systematischen Auswahl geeigneter Aufwandschätzmethoden zu entwickeln
- Die Praktikabilität des Konzepts anhand eines Beispiels zu illustrieren
- Die Integration verschiedener Methoden aufzuzeigen (am Beispiel Case-based Reasoning und Expertenschätzung)

Das Vorgehen zur Methodenauswahl greift auf Elemente der Qualitätssicherungsmethode Quality Function Deployment (QFD)¹ zurück. Ziel von QFD ist es, durch Erfassung lösungsunabhängiger Anforderungen den Raum möglicher Lösungen zu erweitern, und durch die Gewichtung der Anforderungen die Lösungen zu priorisieren. Analog dazu soll die Auswahl einer Schätzmethode nachvollziehbar und auf Anforderungen gegründet sein und der Weg für neue, d. h. bisher ungenutzte Methoden geebnet werden.

Hierzu wird zunächst in Kapitel 2 ein Überblick über existierende Aufwandschätzmethoden, insbesondere deren Einsatzvoraussetzungen, und Verfahren zur Integration verschiedener Methoden gegeben. In Kapitel 3 wird dann der Ansatz zur Methodenauswahl erläutert. Dazu werden Annahmen getroffen, deren Erfüllung für die exemplarisch dargestellte Auswahlmatrix notwendig ist. Anschließend wird die Vorbereitung der Auswahl beschrieben, eine Übersicht über die Anforderungserfüllung einzelner Schätzmethoden gegeben und der Auswahlprozess geschildert. In Kapitel 4 wird das Vorgehen anhand eines Anwendungsbeispiels illustriert. Schließlich erfolgt in Kapitel 5 eine Diskussion der Ergebnisse.

2 Stand der Forschung

Aufwandschätzung von Softwareprojekten ist Untersuchungsgegenstand mehrerer Wissensgebiete: Softwareentwicklung, Forecasting und Projektmanagement. Fokus dieses Beitrags ist die Anwendung in der Softwareentwicklung.

2.1 Aufwandschätzmethoden in der Softwareentwicklung

Aufwandschätzmethoden können grob unterteilt werden in solche basierend auf Statistik (parametrische Modelle), Erfahrung bzw. Erfahrungsaufbau (Expertenschätzung) und Machine learning [BESW98, 2]. Zur weiteren Unterteilung finden sich unterschiedliche Ausführungen (z. B. gemäß der erforderlichen Menge historischer Projektdaten; zum Überblick vgl. [Molø04, 36]). Zur Erläuterung des Auswahlansatzes werden für jede Obergruppe Beispiele genannt² und anhand der Einsatzvoraussetzungen charakterisiert.³

¹ zu QFD siehe z. B. [Akao92; HeSM97]

² Nicht behandelt werden Methoden, die zwar als solche bezeichnet werden, aber streng genommen keine Schätzmethoden sind, z. B. „Price to win“ oder „Parkinson’sches Gesetz“ [Kite96, 190].

³ Um das Auswahlvorgehen zu erläutern, bedarf es keiner lückenlosen Methodenübersicht. Da allerdings je nach Situation unterschiedliche Methoden zur Schätzung geeignet sein können, müssen derart unterschiedliche Methoden zur Auswahl berücksichtigt werden.

2.1.1 Statistische Methoden (parametrische Modelle)

Beispiele statistischer Methoden, die zur Modellierung im Rahmen der Aufwandschätzung eingesetzt werden, sind einfache, schrittweise oder multiple Regressionsanalyse [BEPW06], Analysis of Variance (ANOVA, [BLW99]), Classification and Regression Trees (CART, [BESW98]) und Optimal Subset Selection [BoFW74]. Das bekannteste parametrische Modell, das auf statistischen Methoden basiert, ist CoCoMo [Boeh00].

Voraussetzung für den Einsatz sind Kenntnisse oder begründete Vermutung über die Kausalzusammenhänge zwischen Einflüssen und Aufwand [BEPW06, 8]. Alle Parameter des Modells müssen in numerischer Form vorliegen oder in eine solche transformiert werden. Außerdem muss der Umfang der Datenbasis ermöglichen, statistisch signifikante Korrelationen zwischen Einflussfaktoren und Aufwand bzw. zwischen Einflussfaktoren untereinander zu bestimmen. Die Daten müssen je nach statistischem Verfahren unterschiedliche Eigenschaften aufweisen⁴ (z. B. standardnormal verteilt sein, keine Kollinearität besitzen, vgl. [BEPW06]).

2.1.2 Expertenschätzung

Die Schätzung durch einen oder eine Gruppe von Experten kann formalisiert (z. B. mittels Analytical Hierarchy Process oder Delphi-Methode) oder ad hoc und informell ablaufen. Ziel der Schätzung durch eine Gruppe ist es, durch Austausch und gegenseitiges Feedback unter den Teilnehmern einen Konsens bzgl. des Schätzwertes zu erreichen, in den die Erfahrung und jeweils berücksichtigte Einflüsse aller Teilnehmer eingehen. Voraussetzung des Einsatzes dieser Methoden ist die Erfahrung der Beteiligten mit den Aufgaben⁵, für die sie den Aufwand schätzen (für weitere Details zu Expertenschätzung s. z. B. [Jørg05]).

2.1.3 Machine learning

Hierzu zählen alle Methoden, die Fähigkeiten des menschlichen Verstands nachbilden [MKLP00]. Sie eignen sich insbesondere zur Lösung sehr komplexer Probleme und vermögen aus Erfahrung zu lernen [AaP194, 4; BEPW06, 14]. Bekanntestes Beispiel sind künstliche Neuronale Netze [Lang04]. Weitere Vertreter dieser Methodengruppe sind Case-based Reasoning [AaP194], Genetic Programming [Dola01] und Rule Induction [MKLP00]. Voraussetzungen für den Einsatz dieser Methoden sind eine numerisch vorliegende Datenbasis und die Konfigura-

⁴ Liegen die Eigenschaften bzw. Voraussetzungen nicht vor, ist die Anwendung des jeweiligen statistischen Verfahrens nicht zulässig (s. z. B. [BEPW06], S. 78 und S. 430)

⁵ Erfahrung wird hier verstanden als Durchführung einer Aktivität (bspw. Test) in mehreren vorangegangenen Projekten.

tion des jeweiligen Machine learning-Systems, d. h. Treffen einer Reihe von Aufbau- bzw. Design-Entscheidungen [BEPW, 763; KCCS00, 4].

2.2 Integration verschiedener Methoden

Zahlreiche Autoren kommen zum Schluss, dass in jedem Schätzvorgang verschiedene Methoden eingesetzt werden sollen: „Almost any combination of forecasts proves more accurate than the single inputs“ [BIHo90, 889; s. auch BuFa04, 33; Devn99; Good05; Kite96; Jørg05; WaJe98]. Vielfach wird empfohlen, die Ergebnisse verschiedener Methoden zu vergleichen, ohne jedoch anzugeben, wie im Falle abweichender Ergebnisse vorzugehen ist [WaJe98, 25]. Kitchenham nennt dieses Vorgehen „validation by alternative estimates“ [Kite96, 190]. Um Stärken verschiedener Methoden zu vereinen, müssen sie jedoch integriert werden anstatt die Ergebnisse zur Validierung zu vergleichen.

Integrationsansätze unter Beteiligung von Expertenschätzung lassen sich unterscheiden in mechanische / festgelegte und interaktive / freiwillige Kombination der Methoden [Good00].

Autoren, die *mechanische Kombination von Methoden* untersuchen, wenden mathematische Verknüpfungen der gleichzeitig gewonnenen Schätzergebnisse an. Blattberg und Hoch bspw. gewichteten Expertenschätzung und Ergebnisse statistischer Modelle gleich und addierten jeweils 50 % des Schätzwerts [BIHo90]. Boehm et al. integrierten Expertenschätzung und Regressionsanalyse bei der Entwicklung des Modells CoCoMo [Boeh00], indem eine Gruppe von Experten initiale Skalenwerte für die Einflussfaktoren des Modells festlegte. In [Devn99] wird beschrieben, wie mittels Anwendung des Bayes-Theorems Expertenschätzung und statistische Analyseergebnisse verknüpft und damit CoCoMo kalibriert wurden [Devn99, 37ff].

Diese Ansätze sind entweder pragmatischer Natur [BIHo90], d. h. es wurde nicht untersucht, ob eine andere Integrationsweise zu genaueren Schätzergebnissen führt. Oder sie sind mit hohem Kalibrierungsaufwand (zur Anpassung an Gegebenheiten der schätzenden Organisation) verbunden, damit ausreichend genaue Schätzergebnisse resultieren (s. z. B. [Kite96]). Vorteile der mechanischen Integration, die vor Verfügbarkeit der Schätzergebnisse festgelegt wird, sind die objektive, nachvollziehbare Vorgehensweise, die Manipulation der Ergebnisse verhindert, und die Verminderung der Expertenschätzung inhärenten Inkonsistenz [Good05].

Interaktive Methodenintegration überlässt dem Schätzenden die Gewichtung des statistisch (oder durch ein Machine learning-System) ermittelten Schätzwerts [Good00, 86]. Bei der Schätzung kann der Experte entscheiden, ob und wie er das vorliegende Schätzergebnis verwendet.

Beispielsweise kann der von einem parametrischen Modell ermittelte Schätzwert von einem Experten genutzt werden, seine auf Analogieschluss basierende Schätzung zu korrigieren, indem er die Grenzwerte seines geschätzten Intervalls reduziert.

Nachteil der interaktiven Integration ist, dass Expertenschätzung durchgängig von übermäßigem Selbstvertrauen geprägt ist [GoWr04, 371]. Auch andere systematische Fehler menschlicher Wahrnehmung und menschlichen Urteilens [z. B. BoDö02, 182ff] können zum Tragen kommen [BIHo90]. Schließlich unterliegen Experten politischen Einflüssen innerhalb der Organisation und kombinieren Beobachtungen nicht konsistent [BIHo90, 889f]. Vorteil dieser Integrationsweise ist, dass sie eher akzeptiert wird, weil der Schätzende die Kontrolle über das endgültige Ergebnis der Schätzung behält [BIHo90]. Er kann zudem Veränderungen des Umfelds und besondere Ereignisse berücksichtigen, die aufgrund ihrer Seltenheit nicht modelliert werden. Außerdem hat er die Möglichkeit, bei Aufteilung der Schätzaufgabe zwischen Teilen zu unterscheiden, die durch ein Modell geschätzt werden können und solchen, die mittels Intuition besser schätzbar sind [Good00, 88]. Schließlich kann Schätzung durch einen Experten fehlende oder nicht verlässliche historische Daten kompensieren [Good05, 9].

3 Prozess zur systematischen Methodenauswahl für die Aufwandschätzung

Für die hier präsentierten Zusammenhänge zwischen Anforderungen und Schätzmethoden werden folgende Annahmen getroffen⁶:

Annahme 1: Daten für Modellierung mit statistischen Methoden oder zum Training von Machine learning-Systemen sind vorhanden.

Das bedeutet, dass für mehrere abgeschlossene Projekte verlässliche, vergleichbare Daten, möglichst verhältnisskaliert oder intervallskaliert, verfügbar sind.⁷

Ist Annahme 1 nicht erfüllt, können parametrische Modelle bzw. Machine learning-Systeme nur mit organisationsfremden Daten erstellt bzw. trainiert werden. Dadurch sind sie angesichts

⁶ Perspektive ist die einer softwareentwickelnden Organisation, die verschiedene Produkte als Standardsoftware weiterentwickelt, gelegentlich technisch überarbeitet und regelmäßig neue Produkte entwickelt, und in der bislang ohne Methode Aufwand geschätzt wird.

⁷ Die erforderliche Anzahl der Projekte richtet sich nach der Anzahl der unabhängigen Variablen, die modelliert werden sollen, d. h. nach der abzubildenden Komplexität. Je mehr unabhängige Variablen zur Erklärung der abhängigen Variablen, also des zu schätzenden Aufwands, heran gezogen werden sollen, umso mehr Projekte müssen analysiert werden.

bestimmter Anforderungen weniger zum Einsatz geeignet (z. B. sind Akzeptanz und Schätzgenauigkeit in geringerem Ausmaß gewährleistet [Kitc96]).

Annahme 2: In der Organisation wird zum Zeitpunkt der Auswahl keine formalisierte Methode eingesetzt.

Diese Annahme besagt, dass z. B. keine Kalibrierung des algorithmischen Schätzmodells CoCoMo stattgefunden hat, und dass kein Machine learning-System konfiguriert ist.

Ist Annahme 2 nicht erfüllt, weisen bereits im Einsatz befindliche formalisierte Methoden andere Werte bezüglich der Anforderungserfüllung auf (insbesondere bezogen auf die Anforderung „Ergebnis aufwandsarm verfügbar“).

Sind Annahme 1 oder 2 nicht erfüllt, müssen die Erfüllungsgrade des Templates (s. Tabelle 1) neu bestimmt werden (für bereits eingesetzte Methoden).

3.1 Vorbereitung der Methodenauswahl

Im ersten Schritt werden die Ziele der Schätzung festgehalten. Unter anderem ist zu ermitteln, für wen und welchen Zweck Aufwand geschätzt werden soll.

Im zweiten Schritt werden verfügbare Zeit, Ressourcen und Budget für die Schätzung festgelegt. Die Ergebnisse dieser zwei Schritte, Ziele und Rahmenbedingungen der Schätzung, repräsentieren wesentliche Anforderungen.

Anschließend müssen die Stakeholder identifiziert werden, deren Interessen bei der Auswahl zu berücksichtigen sind (z. B. Budgetverantwortliche, oberes Management). Schließlich werden die Anforderungen⁸ aller Stakeholder an die Schätzmethode und verfügbare Methoden gesammelt bzw. die im Template vorhandenen Listen überarbeitet (s. Tabelle 1). Die Anforderungen werden in den Zeilen, die Schätzmethoden in den Spalten eingetragen⁹.

3.2 Anforderungserfüllung der Schätzmethoden

Jede Schätzmethode trägt zur Erfüllung jeder Anforderung abhängig von ihren Stärken und Schwächen in unterschiedlichem Ausmaß bei. Die Skala der Anforderungserfüllung

⁸ Die Anforderungsermittlung ist für die Methodenauswahl entscheidend. Zum Vorgehen s. [HeSM97], S. 80ff.

⁹ Dies folgt dem Aufbau des House of Quality, in dem Anforderungen den Lösungsmerkmalen gegenübergestellt werden, um die Erfüllung der Anforderungen je Lösungsmerkmal zu untersuchen, vgl. [HeSM97], S. 43f.

(Erfüllungsgrade) reicht von keinem Beitrag zur Erfüllung (mit Wert 0 versehen) über geringe (1) und mittlere Erfüllung (3) bis zu sehr hoher Erfüllung der Anforderung (Wert 9).¹⁰

Tabelle 1 gibt die analysierten Erfüllungsgrade wieder und kann als Template verwendet werden. Enthält eine Zelle mehrere Erfüllungsgrade, kommt es auf das Vorliegen bestimmter Bedingungen an, welcher Wert für eine konkrete Anwendungssituation zutreffend ist. Die notwendigen Bedingungen sind mit Klein- und Großbuchstaben gekennzeichnet und in der Legende erklärt. Die Begründung der ermittelten Erfüllungsgrade wird im Anhang beispielhaft anhand der Anforderung „Ergebnis aufwandsarm verfügbar“ erläutert.

Anforderungen: Ziele u. Rahmenbedingungen d. Schätzung	Schätzmethode	Schätzmethode													
		parametr. Modell (Regr. analyse)	parametr. Modell (optimal subset sel.)	CoCoMo	Classification and Regression Trees	Analogieschluss Informell	indiv. Expertenschätzung: ad hoc	indiv. Expertenschätzung: formalisiert	Gruppe von Experten: ad hoc	Gruppe von Experten: formalisiert	künstl. Neuronale Netze	Rule Induction	Genetic Programming	Case-based reasoning	
Ergebnis aufwandsarm verfügbar	0%	1	1	3	3	9	9	9	3	3	1	1	3	3	
Nachvollziehbar	0%	9	9	9	9	9	9	1	3	1	3	0	9	1	9
Hohe Schätzgenauigkeit und Treffsicherheit	0%	0 d,ST 1 D,ST 3 D,MST	1 M,ST 3 D,ST 9 D,MST	0 d,ST 1 D,ST 3 D,MST	1 d,ST 3 D,ST 9 D,MST	1 k 3 A 9 A	1 kSE 3 RSE 9 Kstat,R,SE	1 3 R 9 Kstat,SE,VE	3 9 R 9 Kstat,SE,VE	3 R 9 R,SE,VE	1 3 D,V 9 D,ST,V	1 3 D,V 9 D,ST,V	1 3 D,V 9 D,ST,V	1 k 3 A 9 A,kex	
Akzeptanz durch oberes Management	0%	9	9	9	1	3	1	3	3	3	1	1	1	1	9
Akzeptanz durch Projektmitarbeiter	0%	3	3	3	3	3	3	3	9	9	1	1	1	1	9
Ohne Verwendung historischer Daten	0%	0	0	1	0	0	0	9	9	9	0	0	0	0	9
Kein Domänenwissen erforderlich	0%	3	3	9	3	1	0	0	0	0	9	3	9	3	9
leicht anpassbar (an neue Technologie, Entw.prozess, ...)	0%	1	1	1	1	3	9	9	9	9	1	1	1	1	1
Ohne Zusatzkosten f. Softwarelizenz	0%	1	0	1	0	9	9	9	9	9	0	0	0	0	0
Kein statistisches Wissen erforderlich	0%	0	0	1	1	9	9	9	9	9	9	9	9	9	9
Automatisierbar	0%	9	9	9	9	3	0	1	0	1	9	9	9	9	9
Früh im Entwicklungsprozess verwendbar	0%	3 M	3 M	3 M	3	9	9	9	9	9	3	3	3	3	9
Leicht manipulierbar	0%	1	0	0	3	9	9	3	9	3	1	3	3	3	9
Objektiv (intersubjektiv durchführbar)	0%	9	9	9	9	0	0	0	1	3	9	9	9	3	9
Jederzeit wiederholbar m. möglicher Verbesserung	0%	1 M	1	1	1 M	9	9	9	9	9	1 M	1 M	1 M	1 M	9
Berücksichtigung einer Vielzahl v. Aufwandseinflüssen	0%	3	3	9	3	9	9	9	9	9	3	1	9	9	9
Unabhängig v. Test- o. Entwicklungsmethodik, -prozess, ...	0%	0	0	0	9	9	3 EU	3 EU	3 EU	3 EU	9	9	9	9	9
...															
Absolute Bedeutung															
Relative Bedeutung															
Rangfolge															

LEGENDE

0/1/3/9 keine / geringe / mittlere / sehr hohe Erfüllung
 ä / A / ka mind. ein ähnliches / sehr ähnliches abgeschlossenes Projekte existiert / kein ähnliches Projekt existiert
 D / d hohe / niedrige Datenqualität (hoch Präzision stat. Methoden erfüllt, große Anz. Beobachtungen, keine fehl. Werte, verlässl. Daten)
 dok Vorgehen und Ergebnisse verständlich dokumentiert
 EU Erfahrung mit Entw.umgebung (Technologie, Entw.prozess, fachl. Anforderungen) liegt vor
 Kstat / Kex Kombination mit stat. Methode / mit Expertenschätzung
 M / m schätzrelevante Merkmalsausprägungen des zu schätzenden Projekts frühzeitig (Planungsphase) bekannt / nicht bekannt
 R Risikoanalyse ist Teil d. Schätzung
 SE / kSE Schätzerfahrung liegt vor / nicht vor
 ST / st hohe / geringe Stabilität des Entw.umfelds (Technologie, Entw.prozess, fachl. Anforderungen)
 V Vielzahl von Einflussfaktoren erfasst
 vE verschiedene Erfahrungsintergründe liegen vor

Tabelle 1: Anforderungen (ungewichtet), Schätzmethode und Erfüllungsgrade

3.3 Vorgehen der Methodenauswahl

Die folgenden Schritte sind jeweils zur Auswahl geeigneter Methoden zu durchlaufen. Basis ist das Template (s. Tabelle 1), das jeweils ergänzt und überarbeitet wird.

1. Auswahl bzw. Ergänzung und Gewichtung der Anforderungen

¹⁰ Über die vier Bewertungsstufen keine, schwache, mittelmäßige und starke Anforderungserfüllung hinaus sind feinere Unterteilungen nur angemessen, wenn sie gerechtfertigt werden können. Gemäß allgemeiner Konvention in der Anwendung von QFD werden diesen Erfüllungsgraden die Werte 0, 1, 3 und 9 zugeordnet [HeSM97, 115].

Die in der jeweiligen Schätzsituation relevanten Anforderungen (neben Zielen und Rahmenbedingungen) müssen identifiziert und anschließend unter Berücksichtigung aller Stakeholder (s. Kapitel 3.1) gewichtet werden.¹¹ Die Gewichte werden neben den Anforderungen im Template erfasst.

2. Ergänzung und gegebenenfalls Auswahl verfügbarer Schätzmethoden

Es wird überprüft, ob die vorliegende Sammlung um (neue) Schätzverfahren ergänzt werden muss. Sind bestimmte Schätzmethoden nicht verfügbar (z. B. statistische Methoden, wenn keinerlei statistisches Wissen vorliegt oder kurzfristig erworben werden kann), können sie von der Auswahl ausgeschlossen werden.

3. Ermittlung der Bedeutung (absolut und relativ) und Rangfolge der Schätzmethoden

Mittels Gewichtung der Anforderungen und Ausmaß der Anforderungserfüllung lassen sich die absolute Bedeutung, die relative Bedeutung und davon abgeleitet die Rangfolge der Schätzmethoden ermitteln. Dazu werden für jede Schätzmethode (in Spalten in Tab. 1) die Erfüllungsgrade mit dem relativen Gewicht der jeweiligen Anforderung multipliziert und aufaddiert. Das Ergebnis ist die absolute Bedeutung der Schätzmethode. Die relative Bedeutung ergibt sich nach Division jedes einzelnen Absolutwerts durch die Summe der absoluten Bedeutungen. Entsprechend den relativen Bedeutungen wird die Rangfolge der Schätzmethoden ermittelt.¹²

4. Auswahl der Schätzmethoden

Die systematische Auswahl der Schätzmethoden folgt deren Rangfolge. Allerdings wird nicht eine Methode gewählt, sondern die zwei oder drei ranghöchsten Methoden. Die Anzahl hängt vom verfügbaren Wissen zur jeweiligen Methode und von der Basis der Methoden ab (Statistik, Expertenschätzung, Machine learning). Sie sollten sich stark unterscheiden, damit Zusatznutzen von mehrfacher Schätzung entstehen kann. Dies folgt den Ergebnissen der Mehrzahl wissenschaftlicher Arbeiten, nach denen eine Kombination verschiedener Methoden einzelnen Methoden überlegen ist (vgl. Kapitel 2.2).

Ohne Kenntnisse einer Methode erreichen damit gewonnene Schätzergebnisse deutlich geringere Treffsicherheit und Genauigkeit, wie experimentell belegt wurde [WaJe98, 26].

¹¹ Um die Gewichtung bei sehr vielen Anforderungen zu vereinfachen, werden die Anforderungen zuerst gruppiert und die Gruppen mit Gewichten versehen. Im nächsten Schritt werden die Anforderungen jeder einzelnen Gruppe gewichtet. Multipliziert man die Einzelgewichtungen mit dem jeweiligen Gruppengewicht, erhält man die absolute Gewichtung jeder Anforderung. Um relative Gewichte zu erhalten, wird jedes absolute Gewicht durch die Summe aller absoluten Gewichte dividiert.

¹² Ein anschauliches Beispiel zur Ermittlung der Bedeutungen findet sich bei [HeSM97, 41f].

Unterschiedliche Fundamente der ausgewählten Methoden sind wichtig, um die Validierung der Ergebnisse zu ermöglichen [Kitc96]. Außerdem können nur unter dieser Voraussetzung die Stärken der einzelnen Methoden bei gleichzeitiger Vermeidung der Schwächen genutzt werden.

4 Anwendungsbeispiel

Das beschriebene Beispiel ist fiktiv, jedoch so realitätsnah wie möglich gewählt. Die Form der Fallstudie zur Überprüfung des Auswahlansatzes eignet sich, weil das „Wie“ der Schätzung interessiert, ohne dass der Forscher die Vielzahl von Einflussfaktoren messen oder steuern kann, denen der Untersuchungsgegenstand (die Schätzmethodenauswahl) unterliegt [Yin03, 8f].

In einem softwareentwickelnden Unternehmen soll die Aufwandschätzung methodisch untermauert werden, um nachvollziehbare, verlässliche Schätzungen zu ermöglichen. Ein Mitarbeiter der Qualitätssicherungsabteilung wird mit der Vorbereitung der Auswahl betraut.

Als Stakeholder werden die Projektleiter identifiziert, die für die Aufwandschätzung verantwortlich sind, Entwickler und Tester, für die die Schätzergebnisse Charakter von Zielvorgaben haben, die Managementebene, die Projektbudgets genehmigt, sowie die Qualitätssicherungsabteilung, die für den Entwicklungsprozess verantwortlich ist und die Kosten der Auswahl trägt.

Mehrere Vertreter jeder dieser Gruppen werden per E-Mail oder telefonisch nach ihren Anforderungen an eine Schätzmethode gefragt. Dazu wird ihnen die Liste von Anforderungen des Templates gegeben, die sie ergänzen bzw. reduzieren können.

Um die Anforderungen zu konsolidieren und zu gewichten, wird je ein Vertreter zu einer Sitzung eingeladen. Nicht teilnehmende, vorher Befragte erhalten die Ergebnisse zugeschickt und werden gebeten, begründete Einwände mitzuteilen. Aufgrund mehrerer Einsprüche werden die ursprünglich geplanten Teilnehmer zu einer zweiten Sitzung eingeladen. Jeder Teilnehmer (drei Projektleiter, zwei Mitarbeiter der Qualitätssicherung, ein Vertreter des Managements mit Verantwortung für Budgetplanung) erhält gleiches Gewicht der Stimme zur Anforderungsgewichtung. Die resultierende Liste gewichteter Anforderungen ist in Tabelle 2 dargestellt.

Die im Template enthaltenen Erfüllungsgrade werden mit einem externen Aufwandschätzexperten überprüft. Insbesondere die Erfüllungsgrade der Anforderung „Hohe Schätzgenauigkeit und Treffsicherheit“ werden entsprechend der Situation, in der sich die Mehrzahl der Projekte befindet, angepasst. Das Ergebnis ist ebenfalls in Tabelle 2 abgebildet.

Anforderungen: Ziele und Rahmenbedingungen der Schätzung	Schätzmethoden	Schätzmethoden												
		parametrisches Modell basierend auf schrittweiser R	parametrisches Modell basierend auf optimal subset	CoCoMo	CART (Classification and Regression Trees)	Analogieschluss informell	individuelle Expertenschätzung: ad hoc	individuelle Expertenschätzung: formalisiertes Vorgehen	Gruppe von Experten: ad hoc	Gruppe von Experten: formalisiertes Vorgehen	künstl. Neuronale Netze	Rule Induction	Genetic Programming	Case-based reasoning
Ergebnis aufwandsarm verfügbar	6%	1	1	3	3	9	9	9	9	3	1	1	3	3
Nachvollziehbar	3%	9	9	9	9	9	1	3	1	3	0	9	1	9
Hohe Schätzgenauigkeit und Treffsicherheit	15%	1	3	1	3	3	1	3	9	9	1	1	1	3
Akzeptanz durch oberes Management	15%	9	9	9	1	3	1	3	3	3	3	3	3	9
Akzeptanz durch Projektmitarbeiter (Entwickler, Tester,...) -> akzeptierte	15%	3	3	3	3	3	3	3	9	9	1	1	1	9
Ohne Verwendung historischer Daten	0%	0	0	1	0	0	9	9	9	9	0	0	0	9
Kein Domänenwissen erforderlich	0%	3	3	9	3	1	0	0	0	0	9	3	9	3
leicht anpassbar (an neue Technologie, Entw.umgebung, Entw.prozess, fachl. Hintergrund,...)	8%	1	1	1	1	3	9	9	9	9	1	1	1	1
Ohne Zusatzkosten f. Softwarelizenz	0%	1	0	1	0	9	9	9	9	9	0	0	0	0
Kein statistisches Wissen erforderlich	0%	0	0	1	1	9	9	9	9	9	9	9	9	9
Automatisierbar	6%	9	9	9	9	3	0	1	0	1	9	9	9	9
Früh im Entwicklungsprozess verwendbar	20%	3	3	1	3	9	9	9	9	9	3	3	3	9
Leicht manipulierbar	0%	1	0	0	3	9	9	3	9	3	1	3	3	9
Objektiv (verschiedene Personen müssen zum gleichen Ergebnis kommen)	3%	9	9	9	9	0	0	1	3	9	9	9	9	3
Wiederholbar mit möglicher Verbesserung (jederzeit im	5%	1	1	1	1	9	9	9	9	9	1	1	1	9
Berücksichtigung einer Vielzahl von Aufwandseinflüssen	5%	3	3	9	3	9	9	9	9	9	9	3	3	9
Unabhängig einsetzbar (von Test- oder Entwicklungsmethodik, Test- oder Entwicklungsprozess, Entwicklungssprache, Entwicklungsarchitektur,...)	0%	0	0	0	9	9	3	3	3	3	9	9	9	9
Absolute Bedeutung		3,9	4,2	3,9	3,1	5,2	4,7	5,4	7,2	7,0	2,3	2,2	2,2	7,0
Relative Bedeutung		6,7	7,2	6,7	5,3	9,0	8,1	9,4	12,3	12,0	4,0	3,8	3,7	11,9
Rangfolge		9	7	8	10	5	6	4	1	2	11	12	13	3

Tabelle 2: Anwendungsbeispiel: gewichtete Anforderungen, angepasste Erfüllungsgrade

Da „Schätzung durch eine Expertengruppe ad hoc“ und „Schätzung durch eine Expertengruppe: formalisiert“ das gleiche Fundament aufweisen, ist wenig Zusatznutzen aus dem Einsatz beider Methoden zu erwarten. Eine Differenzanalyse wird durchgeführt, um die Methode zu identifizieren, die wichtigeren Anforderungen möglichst deutlich besser erfüllt.

Die Methoden „Schätzung durch eine Expertengruppe ad hoc“ und „Case-based Reasoning“ werden ausgewählt. Zum einen ist bei Betrachtung der relativen Bedeutung ein deutlicher Abstand zur rangnächsten Methode erkennbar. Zum anderen weisen die ausgewählten Methoden komplementäre Fundamente auf (Expertenschätzung und Machine learning), weshalb sie sich zur Integration eignen (zu dieser speziellen Kombination siehe [Kitc96]).

Die Integration der Methoden erfolgt interaktiv. Dies liegt zum einen an mangelnder Erfahrung mit Case-based Reasoning, zum anderen an den häufigen Änderungen und Unvorhersehbarkeit des Entwicklungsumfelds.

Die Konfiguration des Case-based Reasoning-Tools setzt Erfahrung mit Aufwandschätzung voraus. Der Qualitätssicherungsmitarbeiter wird hierbei von zwei erfahrenen Projektleitern

unterstützt. Gemeinsam treffen sie Entscheidungen zur Auswahl der Projektmerkmale, Skalierung der Merkmale, Bestimmung des Ähnlichkeitsmaßes, Anzahl zu verwendender Analogieobjekte und Anpassung des Aufwands der Analogieobjekte [KCCS00, 4].

Mit dem konfigurierten Case-based Reasoning-Tool wird in zwei Pilotprojekten der Aufwand geschätzt. Für das eine finden sich drei sehr ähnliche Vergleichsprojekte, deren Aufwand nur gering streut. Für das andere Pilotprojekt, dessen Umfeld von wesentlich größerer Unsicherheit geprägt ist, lässt sich nur ein Analogieprojekt finden, das sich in einigen Merkmalen stark vom zu schätzenden Pilotprojekt unterscheidet. Die Ergebnisse – ein relativ schmales und ein breites Aufwandsintervall – werden jeweils einer Gruppe von Experten vorgelegt, bestehend aus Projektleiter, einem Entwickler, Designarchitekten und dem Qualitätsverantwortlichen des Projekts. Die Experten nutzen das Ergebnis als Basis ihrer ad hoc-Schätzung, die in einer kurzen Besprechung diskutiert und gemeinsam bestimmt wird. Im Fall des schmalen Aufwandsintervalls wird die Intervallbreite beibehalten und nach oben korrigiert, im anderen Fall wird ein weniger breites Intervall am unteren Rand des ursprünglichen Werts geschätzt.

Insgesamt erfordert das Vorgehen zur Auswahl 50 Arbeitsstunden für Identifikation der Stakeholder und Auswahl der Teilnehmer, Versenden und Auswertung der Anforderungsbefragung, zwei Sitzungen zur Festlegung und Gewichtung der Anforderungen, Überarbeitung des Templates und Ermittlung der auszuwählenden Methoden.

5 Diskussion und Ausblick

Der Aufwand zur Auswahl geeigneter Schätzmethode mag vor allem dem Praktiker hoch erscheinen. Außerdem sind verschiedene Experten am Prozess zu beteiligen, deren Verfügbarkeit sichergestellt werden muss. Und schließlich kann trotz aller Sorgfalt bei der Methodenauswahl jeder Schätzvorgang kein verlässlicheres Ergebnis liefern als die Verlässlichkeit jedes seiner Teilprozesse, und wird „so unzuverlässig wie der schwächste Teilprozess“ sein [BuFa04, 31]. Die beschriebene, formalisierte Vorgehensweise weist allerdings der ad hoc-Auswahl gegenüber einige Vorteile auf:

- Beteiligung aller Stakeholder schafft Verantwortungsgefühl und erhöht die Akzeptanz der Ergebnisse.
- Das Bewusstsein für die Bedeutung der Aufwandsschätzung wird geschärft.

- Kosten für ungeeignete Schätzverfahren werden gespart (z. B. Lizenzgebühren für Software zur Entwicklung künstlicher Neuronaler Netze).
- Gemeinsames Verständnis dessen kann erzielt werden, was ein Schätzverfahren leisten kann (z. B. an Genauigkeit) und welche Voraussetzungen dafür vorliegen müssen.

Ist die Auswahl geeigneter Schätzmethoden für ein Projekt getroffen, kann das Ergebnis in der gleichen Organisation mit geringem Überarbeitungsaufwand (bzw. unter gleichen Bedingungen ohne Anpassung) erneut verwendet werden. Unterstellt man außerdem eine Lernkurve den Auswahlprozess betreffend, kann die zweite Auswahl deutlich schneller ablaufen.

Die praktische Überprüfung des Ansatzes steht noch aus, ist jedoch schon in Vorbereitung. Mit den Ergebnissen der Aufwandschätzung, die unter Verwendung der ausgewählten Methoden stattfindet, soll untersucht werden, ob und inwieweit die systematisch und begründet ausgewählten Methoden zur Verbesserung des Schätzprozesses beitragen können.

Fundament realistischer Schätzung sind historische Daten, deren Qualität (Korrektheit und Validität) und Angemessenheit [BESW98, 8]. Weitere Grenzen sind der Anwendbarkeit dadurch gesetzt, dass Kenntnisse der verschiedenen Methoden vorliegen und Experten (bezogen auf Entwicklung und Schätzung der einzelnen Entwicklungsaufgaben) verfügbar sein müssen.

Wenn der vorgestellte Ansatz neben der Auswahl angemessene Datenerfassung und Projektnachbereitung unterstützt, kann die Aufwandschätzung vom Risikofaktor zur Risikomanagementtechnik [Boeh89, 117] werden und zum Projekterfolg beitragen.

Literaturverzeichnis

- [AaPl94] Aamodt, A.; Plaza, E.: Case-based reasoning: foundational issues, methodological variations, system approaches. In: AI Communications, 7 (1994) 1, S. 39-59.
- [Akao92] Akao, Yoji: Quality Function Deployment. Moderne Industrie, Landsberg 1992.
- [BEPW06] Backhaus, Klaus; Erichson, Bernd; Plinke, Wulff; Weiber, Rolf (Hrsg.): Multivariate Analysemethoden. 11. Aufl., Springer, Berlin / Heidelberg 2006.
- [BESW98] Briand, Lionel C.; El Emam, Khaled; Surmann, Dagmar; Wieczorek, Isabella; Maxwell, Katrina D.: An Assessment and Comparison of Common Software Cost Estimation Modeling Techniques. Technical Report ISERN-98-27

- [BIHo90] Blattberg, Robert C.; Hoch, Stephen: Database Models and Managerial intuition: 50% Model + 50% Manager. In: Management Science, 36 (1990) 8, S. 887-899.
- [BoDö02] Bortz, Jürgen; Döring, Nicola: Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 3. Aufl., Springer, Heidelberg 2002.
- [Boeh00] Boehm, B. W.; Abts, C.; Horowitz, E.; Madachy, R.; Reifer, D.; Clark, B.K.; Steece, B.; Brown, A.W.; Chulani, S.: Software cost estimation with Cocomo II. Prentice Hall, Upper Saddle River (NJ/USA) 2000.
- [Boeh89] Boehm, Barry W.: Software Risk Management. IEEE Computer Society Press, Washington 1989.
- [BoFW74] Boyce, D.E.; Farhi; Weischedel, R.: Optimal Subset Selection, Multiple Regression, Interdependence and Optimal Network Algorithms. In: Beckmann, M.; Künzi, H.P. (Hrsg.): Lecture Notes in Economics and Mathematical Systems. Operations Research. Vol. 103. Springer-Verlag, Berlin et al. 1974.
- [BuFa04] Bundschuh, Manfred, Fabry, Axel: Aufwandschätzung von IT-Projekten, 2. Aufl., mitp-Verlag, Bonn 2004.
- [Devn99] Devnani-Chulani, Sunita: Bayesian Analysis of Software Cost and Quality Models. Dissertation, University of Southern California, Berkeley 1999.
- [Dola01] Dolado, J.J.: On the problem of the software cost function. In: Information and Software Technology, 43 (2001) 1, S. 61-72.
- [Good00] Goodwin, Paul: Improving the voluntary integration of statistical forecasts and judgment. In: International Journal of Forecasting, 16 (2000), S. 85-99.
- [Good05] Goodwin, Paul: How to integrate Management Judgment with Statistical Forecasts. In: Foresight 1 (2005) 1, S. 8-12.
- [GoWr04] Goodwin, Paul; Wright, George: Decision Analysis for Management Judgment. 3. Aufl., John Wiley & Sons, Chichester 2004.

- [HeSM97] Herzwurm, Georg; Schockert, Sixten; Mellis, Werner: Qualitätssoftware durch Kundenorientierung. Die Methode Quality Function Deployment (QFD): Grundlagen, Praxis und SAP R/3 Fallbeispiel. Vieweg, Braunschweig 1997.
- [Jørg05] Jørgensen, Magne: Practical Guidelines for Expert-Judgment-Based Software Effort Estimation. In: IEEE Software, 22 (2005) 3, S. 57-63
- [KCCS00] Kadoda, Gada; Cartwright, Michelle; Chen, Liguang; Shepperd, Martin: Experiences using case-based reasoning to predict software project effort. ESERG Technical Report No. 00-09, Bournemouth University, 2000.
- [Kitc96] Kitchenham, Barbara: Software Metrics: Measurement for Software Process Improvement. Blackwell Publishers, Cambridge (USA) 1996.
- [Lang04] Lange, Carsten: Neuronale Netze in der wirtschaftswissenschaftlichen Prognose und Modellgenerierung. Physika-Verlag, Heidelberg 2004.
- [MaSh03] MacDonell, Stephen G.; Shepperd, Martin J.: Combining techniques to optimize effort predictions in software project management. In: Journal of Systems and Software, 66 (2003) 2, S. 91–98.
- [MKLP00] Mair, Carolyn; Kadoda, Gada; Lefley, Martin; Phalp, Keith; Schofield, Chris; Shepperd, Martin: An investigation of machine learning based prediction systems. In: Journal of Systems and Software, 53 (2000) 1, S. 23-29.
- [Molø04] Moløkken-Østvold, K: Eff.ort and schedule estimation of software development projects. Ph.D. Thesis, Dep. of Informatics, University of Oslo, 2004.
- [MySS05] Myrtveit, Ingunn; Stensrud, Erik; Shepperd, Martin: Reliability and Validity in Comparative Studies of Software Prediction Models. In: IEEE Transactions on Software Engineering, 31 (2005) 5, S. 380-391
- [WaJe98] Walkerden, Fiona; Jeffery, Ross: An Empirical Study of Analogy-based Software Effort Estimation. Technical Report 98/8, University of New South Wales, Sydney 1998.
- [Yin03] Yin, Robert: Case study research. 3. Aufl., Sage, Thousand Oaks (2003)

6 Anhang

Bestimmung der Erfüllungsgrade am Beispiel „Ergebnis aufwandsarm verfügbar“ (vgl. Tabelle 1 und Tabelle 2):

Methoden basierend auf statistischen Verfahren

Die Entwicklung eines parametrischen Modells basierend auf schrittweiser Regressionsanalyse erfordert verlässliche Daten abgeschlossener Projekte und ist trotz erhältlicher Toolunterstützung mit erheblichem Modellierungsaufwand verbunden (z. B. Auswahl der Gleichungsform, vorausgehende Datenanalyse).¹³ Daher erfüllt schrittweise Regressionsanalyse die Anforderung in sehr geringem Maße (Wert 1).

Die Methode „optimal subset selection“ liefert bessere Modellierungsergebnisse als einfache oder schrittweise Regression [BoFW74, 1], ist aber mit vergleichbarem Aufwand zur Modellierung verbunden und seltener von kommerzieller Statistiksoftware unterstützt, weshalb die Erfüllung der Anforderung „Ergebnis aufwandsarm verfügbar“ ebenfalls gering ist (1).

Das Schätzmodell CoCoMo enthält Standardwerte für Koeffizienten und Konstanten der Schätzgleichung, sodass prinzipiell zur Anwendung nur die Werte der modellierten Variablen des zu schätzenden Projekts gemessen und in die Gleichung eingesetzt werden müssen. Allerdings muss das Modell für genaue Schätzergebnisse mit lokalen Projektdaten kalibriert werden, sodass die Anforderung „Ergebnis aufwandsarm verfügbar“ nur mittelmäßig erfüllt wird (3).

Bei Schätzung mittels eines „Classification and Regression Trees“ gibt es unterstützende Softwaretools, doch fällt erheblicher Aufwand zur Datenerfassung an, da wie für alle statistischen Ansätze mehr Daten vorliegen (und eingegeben werden) müssen, je mehr unabhängige Variablen zur Schätzung verwendet werden (bzw. je mehr Verzweigungen der Baum aufweist). Außerdem muss festgelegt werden, welche Projektmerkmale die zu schätzende Größe beeinflussen und demnach zur Bildung des Regressionsbaumes verwendet werden, und unter welcher Bedingung ein Zweigende erreicht ist. Insgesamt wird die Anforderung „Ergebnis aufwandsarm verfügbar“ nur mittelmäßig erfüllt (3).

Expertenschätzmethoden

Bei Verwendung der Methode „Analogieschluss informell“ muss sich der Experte vergleichbare Projekte in Erinnerung rufen. Weder Vergleichsdaten noch Vorgehen sind zu beschreiben,

¹³ Selbst wenn bereits ein solches Modell mit firmeneigenen Daten entwickelt wurde, muss es regelmäßig überprüft und gegebenenfalls aktualisiert werden (z. B. bei Veränderung des Entwicklungsprozesses). Ein Schätzergebnis ist nur in den seltenen Fällen aufwandsarm verfügbar, in denen das Modell kurz zuvor mit Daten neuer Projekte überprüft und aktualisiert wurde.

weshalb der Zeitaufwand zwar abhängig von der Bearbeitungszeit des Schätzers, jedoch absolut betrachtet gering ist. Der Grad der Anforderungserfüllung ist demnach sehr hoch (9).

Wird als Schätzmethode „Individuelle Expertenschätzung: ad hoc“ gewählt, ist ebenfalls keine Dokumentation erforderlich. Der Schätzaufwand hängt von der Arbeitsgeschwindigkeit des Experten bei dieser Aufgabe ab. Das Ergebnis ist relativ schnell und aufwandsarm verfügbar, sodass die Anforderungserfüllung dieser Methode als sehr hoch beurteilt wird (9).

Läuft die „individuelle Expertenschätzung“ formalisiert ab, muss das Vorgehen dokumentiert und die herangezogenen Informationen angegeben werden. Hierfür werden häufig Fragebögen bzw. Guidelines oder Templates verwendet, sodass der Zeitaufwand zwar höher als bei dem Verfahren ad hoc, aber absolut betrachtet gering ist. Die Methode erfüllt daher die Anforderung „Ergebnis aufwandsarm verfügbar“ in hohem Maß (9).

Bei einer „Expertenschätzung durch eine Gruppe von Experten“ findet auf der Suche nach einem Konsens Diskussion zwischen den Beteiligten statt. Läuft dieser Austausch formalisiert und über mehrere Feedbackrunden ab (wie bei der Delphimethode), ist mit deutlich mehr Aufwand als bei individueller Expertenschätzung oder ad hoc-Schätzung durch eine Expertengruppe zu rechnen. Der Gesamtaufwand nimmt mit dem Grad der Formalisierung und der Anzahl beteiligter Experten zu. Es fällt weder Einarbeitungs- noch Modellierungsaufwand an, sodass der Erfüllungsgrad der Anforderung „Ergebnis aufwandsarm verfügbar“ bei der ad hoc-Gruppenschätzung hoch (9), bei formalisierter Gruppenschätzung mittelmäßig ist (3).

Schätzmethoden basierend auf Machine learning

„Künstliche Neuronale Netze“ müssen für den Einsatz konfiguriert werden, was mit erheblichem Aufwand zur Bildung des Netzes verbunden ist [KCCS00, 2]. Der Erfüllungsgrad der Anforderung „Ergebnis aufwandsarm verfügbar“ ist daher als niedrig einzustufen (1).

Zur Schätzung mittels „Rule Induction“ müssen lediglich die vorhandenen Daten abgeschlossener Projekte eingegeben werden, damit Regeln zur Schätzung weiterer Projekte durch die Rule Induction-Applikation generiert werden. Allerdings erfordern Regeln, die zu hinreichend genauen Schätzergebnissen führen, Sorgfalt bei der Auswahl der Merkmale, anhand derer die Regeln aufgestellt werden [MKLP00, 26]. Da diese Auswahl je nach Anzahl verfügbarer Merkmale sehr zeitaufwändig werden kann [MKLP00, 26], und bei Softwareentwicklungsaufwand sehr viele mögliche beeinflussende Merkmale existieren, wird der Erfüllungsgrad der Anforderung „Ergebnis aufwandsarm verfügbar“ als niedrig eingestuft (1).

Wird mittels „Genetic Programming“ geschätzt, ist das System anhand verschiedener Parameter zu konfigurieren (z. B. Anzahl der gleichungserzeugenden Durchläufe [Dola01, 64]). Um die vorhandenen Daten ausreichend genau abzubilden, muss der Algorithmus mehrmals mit verschiedenen Parameterwerten durchlaufen werden. Aufgrund des Konfigurationsaufwands wird die Anforderung „Ergebnis aufwandsarm verfügbar“ nur mittelmäßig erfüllt (3).

Die Methode „Case-based Reasoning“ erfordert wie alle Machine learning-Systeme die Eingabe zu analysierender Datensätze und Konfiguration des Systems (z. B. Anzahl auszuwählender Vergleichsprojekte, Auswahl von Vergleichskriterien). Hierfür existiert keine etablierte Vorgehensweise, sodass die Konfiguration meist einem Trial and Error-Prozess mit entsprechendem Aufwand folgt [KCCS00, 20]. Da eine kleine Anzahl Cases für den effektiven Einsatz der Methode ausreicht [KCCS00, 2], wird die Anforderung „Ergebnis aufwandsarm verfügbar“ mittelmäßig erfüllt (3).