




Characteristics that affect preference of decision models for asset selection: an industrial questionnaire survey

Emil Alégroth¹  · Tony Gorschek¹ · Kai Petersen¹ · Michael Mattsson¹

Published online: 28 December 2019
© The Author(s) 2020

Abstract

Modern software development relies on a combination of development and re-use of technical asset, e.g., software components, libraries, and APIs. In the past, re-use was mostly conducted with internal assets but today external; open source, customer off-the-shelf (COTS), and assets developed through outsourcing are also common. This access to more asset alternatives presents new challenges regarding what assets to optimally chose and how to make this decision. To support decision-makers, decision theory has been used to develop decision models for asset selection. However, very little industrial data has been presented in literature about the usefulness, or even perceived usefulness, of these models. Additionally, only limited information has been presented about what model characteristics determine practitioner preference toward one model over another. The objective of this work is to evaluate what characteristics of decision models for asset selection determine industrial practitioner preference of a model when given the choice of a decision model of high precision or a model with high speed. An industrial questionnaire survey is performed where a total of 33 practitioners, of varying roles, from 18 companies are tasked to compare two decision models for asset selection. Textual analysis and formal and descriptive statistics are then applied on the survey responses to answer the study's research questions. The study shows that the practitioners had clear preference toward the decision model that emphasized speed over the one that emphasized decision precision. This conclusion was determined to be because one of the models was perceived faster, had lower complexity, was more flexible in use for different decisions, and was more agile on how it could be used in operation, its emphasis on people, its emphasis on "good enough" precision and ability to fail fast if a decision was a failure. Hence, we found seven characteristics that the practitioners considered important for their acceptance of the model. Industrial practitioner preference, which relates to acceptance, of decision models for asset selection is dependent on multiple characteristics that must be considered when developing a model for different types of decisions such as operational day-to-day decisions as well as more critical tactical or strategic decisions. The main contribution of this work are the seven identified characteristics that can serve as industrial requirements for future research on decision models for asset selection.

✉ Emil Alégroth
Emil.Alegroth@Bth.se

Keywords Decision models · Characteristics · Industrial study · Survey · Model comparison

1 Introduction

Software engineering has gone through a paradigm shift from traditional in-house development of all software assets to an agile way-of-working where software is built through re-use of existing, open-source, and purchased software assets (van der Linden et al. 2009; Badampudi et al. 2016). Deciding what assets to use is a challenge but also paramount to optimize the product's competitive advantage in the market and to ensure its longevity. This challenge is identifiable in all domains, including component-based (Ssaed et al. 2012; Vale et al. 2016) and service-oriented development (Huhns and Singh 2005) or a mix of the two (Breivold and Larsson). Additionally, decisions about asset selection are taken on multiple decision levels; daily operational decisions that affect parts of the software, tactical decisions that affect the software's overall quality, and finally strategic decisions that affect the entire direction of the software system (Badampudi et al. 2016; Moe et al. 2012). However, despite the importance of correct decision-making, many assets are selected purely on expert judgement, ad hoc, or rarely following recommendations from literature (Petersen et al. 2018; Li et al. 2006). These recommendations from literature are divided into three main categories: (1) models that support decision-making (Cortellessa et al. 2006; Potena 2007); (2) what criteria to consider for a given decision (Stewart et al. 2005; Li et al. 2006); (3) ways of comparing different assets (Li et al. 2006; Rudzki et al. 2009). The stated purpose of these recommendations is to improve industrial decision-making about asset selection in terms of cost-effort, precision, or some other measure of decision quality. However, despite this clear purpose, decision support is underutilized in practice, implying that there are other factors that influence their usage (Petersen et al. 2017).

The study presented in this paper is part of a research project, named ORION¹ that has exclusively studied decision-making on asset selection. The project has utilized decision theory (Starmer 2000) as a foundation to produce a taxonomy for decision-making as well as decision models and frameworks (Wohlin et al. 2016; Badampudi et al. 2018). The taxonomy, GRADE (Goals, Roles, Assets, Decisions and Environment) (Martens and Teuteberg 2012; Papatheocharous et al. 2015), takes a wider view at decision-making and the many diverse factors that asset selection includes, e.g., asset origin and asset qualities. GRADE thereby provides guidance for what aspects to consider when developing asset selection support artifacts, such as decision models. However, while outlining the necessary contents and functions of a decision-support model to provide decision quality, the taxonomy does not provide guidance what qualities (also denoted quality characteristics, e.g., ease-of-use/complexity or speed/performance) to consider when developing such models. Qualities, as we know from software development, are key to the acceptance, usability, and value of a product (Dave et al. 2018).

The objective of the study is to identify characteristics of decision models that practitioners consider when selecting a model. Knowledge of these characteristics is valuable for decision-model developers since knowledge of, and emphasis on, these characteristics can help improve the models' adoption and use in practice.

¹ORION project (<http://orion-research.se>), a Swedish research project that began in 2015 involving three Swedish universities

Data collection for this work was carried out in an industrial questionnaire survey where two decision models were presented and then compared by the practitioners. The two models are referred to as ORION Alpha (Wohlin et al. 2016; Badampudi et al. 2018) and APOLLOS (A Priority-Oriented Lightweight, Lean, Object Selection). While capturing all key aspects of GRADE, the models have a different application focus. ORION Alpha focuses on producing very precise and objective decisions, which lengthens the time for a decision. In turn, APOLLOS focuses on quicker and more subjective decisions, which lowers decision accuracy.

The research objective, i.e., identification of model characteristics that influence practitioner preference, has been broken down into two research questions:

- **RQ1:** What preference do industrial practitioners have given the choice between an asset selection decision model with high precision and a model with high speed?

The aims of this question are to (1) identify which of the models, given their varying characteristics, the practitioners prefer and (2) relate this preference to the two models based on their characteristics. This evaluation helps provide initial understanding of the practitioners' preferences and what characteristics influence said preferences.

- **RQ2:** What drawbacks do industrial practitioners' perceive with either model given their varying characteristics?

The objective of this question is to get deeper insight into which characteristics are considered more important.

Synthesis of the answers of these two research questions provides the study's main contribution, which is a list of characteristics that practitioners consider important for a decision model to be applicable and useful. Hence, characteristics that can be regarded as requirements from the practitioners that decision models for asset selection should adhere to. As such, using this list of requirements, researchers can fine-tune their suggested models to align them with practitioner perceptions to achieve better quality and industrial impact. An additional contribution of this work is the qualitative feedback and response from the practitioners about decision models for asset origin and asset selection and their use in practice.

The continuation of this paper starts with a review of related literature in the area of decision theory, decision models, and decision models for asset selection in Section 2. The section will also introduce the compared decision models and an example of their application on a fictional scenario based on empirical observations from industrial practice. This is followed by the study's methodology in Section 3 followed by the research results in Section 4. Next, a synthesis of the results is presented in Section 5.1 and finally the paper is summarized and concluded in Section 6.

2 Related literature and research support

Decision-making and decision theory are vastly studied subjects due to their intricacies and many facets (Starmer 2000), as well as their importance in practice. A majority of this work is however focused on descriptive models that aim, through governance, to help practitioners make better decisions. These works include the decision models developed in the ORION project for asset and asset origin selection (Wohlin et al. 2016; Badampudi et al. 2016) but also related works on customer off-the-shelf (COTS) asset selection (Garg 2017; Maiden and Ncube 1998), open-source asset selection (Ven and Verelst 2006; Li et al. 2011), and

outsourcing (Yang and Huang 2000). Further, as presented by Badampudi et al., most of these works focus on selection of a single asset origin, while in industrial practice several origins are often considered in unison (Badampudi et al. 2016).

Additionally, most of these works focus on functional aspects of the assets. However, functionality is not the only driving factor in a decision, e.g., there are inherent differences between different asset origins, and other factors such as the assets' quality attributes (Borg et al. 2018) also play a major role. These additional factors have therefore been studied in research. For instance, it is well known that attributes such as cost, time, and reliability of the asset (Badampudi et al. 2016) are crucial and that removal of subjective decision-making is critical (Strigini 1996; Aurum and Wohlin 2003). Consequently, requirements for assets and how to model these requirements are well studied. However, when it comes to characteristics, practitioner requirements, and quality attributes of the decision models themselves, very little is known.

This lack of information is vital to understand how to make better decision models and tailor them to various contexts and purposes. A central part toward filling this gap in knowledge is to understand how decision-making is done in practice (Petersen et al. 2017) and use said research to elicit requirements for decision-making models. Decision theory (Berger 2013; Einhorn and Hogarth 1981) provides some support and has, as a subject, been studied in all phases of software engineering (Ruhe 2002). However, once more, little of this work is explicitly focused on the characteristics and attributes that practitioners require for a model to actually be applicable or useful in practice. One reason for this lack of information is the lack of empirical studies from industry on the use of decision-making models that can serve as a foundation for continued research (Cusumano 2004; Franke and Buschle 2016). Studies that explore model characteristics, from an industrial perspective, are therefore needed and warranted to understand what is required for a decision model to provide value to the practitioners in their work.

This study partially bridges this gap in knowledge by studying a new perspective of decision models for asset selection, i.e., what characteristics of the models the industrial practitioners use to determine preference and thereby if a model is applicable and useful.

2.1 The compared decision models

The work presented in this paper is based on the GRADE taxonomy (Martens and Teuteberg 2012; Papatheocharous et al. 2015). The taxonomy describes fundamental concepts of how to perform decision-making and is divided into five concepts: Goals, Roles, Assets, Decisions, and Environment, where:

- **Goals** - The overlaying goals that influence a decision.
- **Roles** - The individuals involved in decision-making.
- **Assets** - The software assets or services that require decision-making.
- **Decision** - The method used to evaluate the available options to identify a solution.
- **Environment** - The context prior to when the decision was made.

These concepts were used to guide the development of the decision models ORION Alpha and APOLLOS that are used in this study. Both models are grounded in the same theoretical foundation acquired from decision theory and GRADE (Papatheocharous et al. 2015; Berger 2013; Einhorn and Hogarth 1981).

ORION Alpha, visualized in Fig. 1, is a comprehensive decision model that focuses on high decision precision using previous decisions and context information to support the decision-maker. Previous decision information is stored in one of the three repositories,

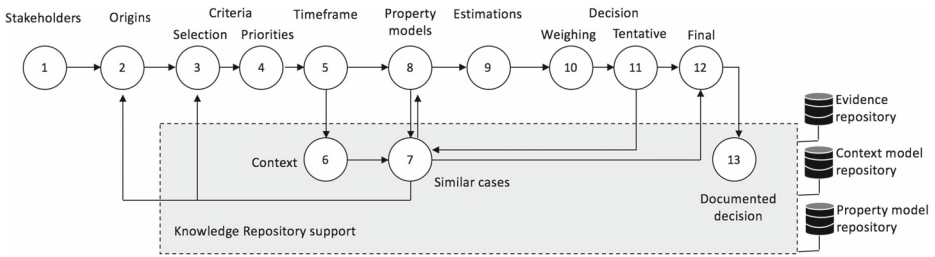


Fig. 1 Visualization of the decision model proposed by Wohlin et al. (2016)

shown in Fig. 1, referred to as decision, property, and context model repositories. The model defines the needs of the stakeholders, what the decision is about, criteria to evaluate the decision and thereby criteria for how to ensure the outcome’s suitability, and decision objectivity. Property models further detail the criteria for comparing assets and methods how to compare them, e.g., COCOMO (Boehm et al. 2000). Finally, a context model is used to capture the circumstances of the decision and to make the decision comparable with future decisions.

The decision-making model also includes a decision process of 13 steps that describes the use of mentioned repositories and other knowledge models. This process is designed for asset origin selection, i.e., if a new asset, when required, should be (1) developed in-house, (2) be outsourced, (3) be acquired as a COTS product, or (4) an open-source library. However, the model is also perceived as usable for asset selection (Wohlin et al. 2016). For more details about ORION Alpha, we refer the reader to the previous published work by Wohlin et al. (2016).

APOLLOS, visualized in Fig. 2, was developed explicitly for this study and is focused on fast decision-making. The model includes five steps:

- **Step 0:** Identification of a problem that requires a decision to be made to find an asset.
- **Step 1:** Identification of potential assets that may solve the problem, which become the basis for the decision.
- **Step 2:** Quick exclusion of assets that do not fulfill minimum requirements, or criteria, to be “good enough” solutions.
- **Step 3:** If a decision turns out to be wrong, the model emphasizes fast failure and new decisions to be taken.

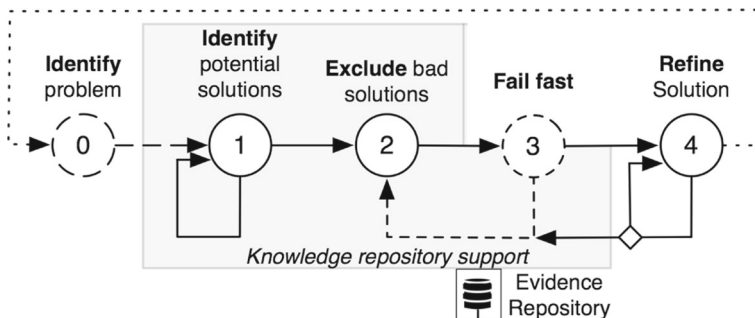


Fig. 2 Visualization of the steps of the APOLLOS framework

- **Step 4:** If a chosen asset is not optimal, it must be refined and may require new decisions.

Step 0 of the model, as shown in Fig. 2, is completed when a specific need for a solution (e.g., a specific technical asset) is identified and delegated to an individual or team to be fulfilled. In this step, it is important to identify the root cause of the problem and distinguish the problem from its symptoms to maximize the positive impact of the final decision. It is also important to identify which stakeholders may be affected by the problem. However, both of these actions are dealt with by the decision-maker who can decide to what level and rigor they should be investigated.

Step 1 encourages the decision-maker(s) to identify as many “good enough” solutions, i.e., technical assets, as possible in an allocated time frame. Factors to consider when seeking solutions are low cost, high reliability, and similarity to previous solutions identified in previous decisions. Similarity can for instance be based on the quality of another asset provided by a specific vendor or open-source community that has been chosen in the past.

Step 2 then aims to minimize the number of possible solutions and quickly reach a single “good enough” solution. “Good enough” is determined subjectively by the decision-maker but should be based on a balance of the cost of introducing an asset compared with the risk and cost of the asset failing to fulfill the identified need for the asset. Minimization of the number of possible assets is also supported by exclusion criteria that the decision-maker(s) decide ad hoc, such as maximum cost, performance needs, and allowed asset origin.

Step 3 is added since the model’s emphasis on speed leaves room for wrong decisions to be taken. The model recognizes that subjective decision-making is associated with additional risk of failure and therefore implores the decision-maker(s) to monitor the success of a decision to identify wrong decisions as quickly as possible. This practice allows the decision to be corrected with perceived minimal wasted time or resources.

Step 4 is required since “good enough” solutions often require refinement to fully solve an identified need. During refinement, it is possible that a solution is found unsuitable and the decision-maker may therefore choose to abandon a solution to take a new quick decision or further refine the solution at additional cost. The decision-maker may also identify cascade problems with the solution that require further decisions to be made. Consequently, the model assumes that software development requires continuous decision-making where one decision leads to new problems that require new decisions and solutions and so forth. Each decision can contribute to an overall risk of decision failure, but the model presents no preventive measure for such escalation.

Finally, the model promotes the use of an evidence repository that can provide the decision-maker(s) with useful information from previously taken decisions. The model does not restrict the nature of the repository, which could thereby be a database, documentation, or expert judgement from an individual.

2.2 Comparison of ORION Alpha and APOLLOS on an industry-inspired case

This section will present a scenario of how ORION Alpha and APOLLOS can be applied in practice. The scenario is inspired by empirical observations made by the authors of the paper at a Scandinavian safety-critical software development company that the researchers have worked with for many years. However, since the scenario did not play out in reality, the company and product will remain nameless. The intent of this comparison is to showcase the two models’ differences and give insight into the time required in each step of each model’s process. Additionally, due to the time spent, and resources used, the scenario aims

to give insight into the precision of each model and thereby the type of decisions the models are useful for, e.g., operational, tactical, and strategic decisions (Badampudi et al. 2016; Moe et al. 2012).

In the scenario, a developer is given the task to integrate an image recognition library into one of the company's products to visually track objects from a camera feed. This functionality is perceived important to the final product and has safety critical use cases which makes the decision tactical for the company. Additionally, due to the feature's importance, stringent control is required over the library to ensure that it can be evolved and the solution must be able to precisely identify moving targets without fail.

To make the comparison of the models as easy as possible for the reader, the decision process has been broken up into five general decision steps: (1) Identify the need for an asset, (2) identify asset alternatives, (3) compare available assets, (4) select an asset, and (5) integrate the asset to solve the need.

Decision Step 1: Identified need for an asset - The starting point for both processes is that there is a need for a new asset with specific functionality and/or purpose, but the processes' entry differ. ORION Alpha assumes that the need for an asset has already been identified before the process begins and does therefore not count the time spent on identifying this need. The identified need may therefore stem from a single developer that has identified that a new component is required in operations. However, the need may also be a strategic decision to move into a new market that requires an asset of a specific type to fulfill some core functionality.

The process then starts with an analysis of all stakeholders that need to be part of the decision before moving on to Decision Step 2. On an operational level, these roles can include the decision-maker (assuming they are developers), their manager, and perhaps some testers. However, on a strategic level, this could be stakeholders scattered across the entire organization.

Regardless of the root cause of the need for a new asset, ORION Alpha relies on an external source, not part of the process, to give the requirements for the asset and to identify what need the asset aims to solve. In the case of the example from the safety critical company, this need was identified by a manager (product owner) on a tactical level that delegated the work to a team of developers.

In contrast, APOLLOS is triggered by the decision-maker when they see a need for a new asset, e.g., an asset to handle dates, a new database, or the image recognition library discussed above. It is then the decision-maker's responsibility to find a minimal set of requirements that need to be fulfilled to satisfy the identified need for a new asset. The process then assumes that the decision-maker is the main stakeholder and it is their responsibility to utilize any resources they see fit, including other stakeholders, to find an asset to fulfill the need. In the case of the safety critical company, this step implied that the responsible developer had to evaluate the needs for the image recognition library and perhaps anchor the acquisition with management.

In comparison with ORION Alpha, this decision step warranted less action from APOLLOS since it does not require the decision-maker to identify or involve other stakeholders if there is no need for it. Further, the decision-maker using APOLLOS is not obligated to analyze the need from any other perspective than their own, making the process more fast moving since it requires less synchronization with other stakeholders.

Decision Step 2: Identify asset alternatives - Any problem has multiple solutions, and in this step, ORION Alpha aims to identify all available solutions. This is done by scanning

multiple asset origins and assessing what assets that are available. Typically, these asset origins include developing a solution in-house, purchasing a COTS, acquiring an open-source asset, or outsourcing the development. Each of these choices has different benefits and drawbacks and shall be analyzed before the actual asset identification can begin. Along with identifying the possible origins, a set of metrics are also defined or developed for the comparison of the available assets. These metrics may be developed for the specific situation but previously used metrics are also collected from a knowledge repository. The metrics are evaluated individually for each possible asset and should include estimation of the longevity of the assets, including available support, maintainability, fit-for-purpose, etc. In the case of the safety critical company, the possible asset origins that were identified were to either develop the functionality in-house at high cost but with high levels of control over the asset, or to use an open-source component or COTS at lower cost but also lower control. Once an asset origin is chosen, the stakeholders associated with the decision must find all assets that seemingly fulfill the requirements and further evaluate them. This evaluation is made using the aforementioned metrics and a predefined assessment procedure that is either re-used from previous decisions or developed for the specific decision. The procedure may also determine a set of specific metrics that it needs for the comparison.

In contrast, APOLLOS, due to the root-cause analysis performed in Decision Step 1, focuses on finding a set of assets that the user thinks are suitable based on expert judgement. For guidance, the user uses historical information of what origins, vendors, and types of assets have been successfully applied in the past. This data is taken from a knowledge repository, but in contrast to ORION Alpha, a repository in APOLLOS case can be a domain expert, e.g., a colleague or manager. For the case of the safety critical company, no real action had to be taken in this step with APOLLOS.

APOLLOS is once more perceived quicker since the model simply suggests the use of more objective metrics but the decision-maker can rely solely on expert judgement. ORION Alpha, however, requires the decision-maker to first analyze how to make the assessment of the assets taking inputs from other stakeholders and analysis methods as well as previous decisions. These differences have clear impact on both time and precision of the decision.

Decision Step 3: Asset comparison - Once all metrics have been collected for all assets, ORION Alpha states that all assets should be compared in regard to all chosen metrics. No consideration is taken to the informative value of each metric and all metrics should be pairwise compared for all possible assets. Once all assets are compared with all metrics, the process proceeds to Decision Step 4. In the case of the safety critical company, this step would entail gathering up developers, managers, and testers to discuss the asset alternatives and/or for the stakeholders to give inputs on the possible assets and get feedback/measurements on each metric.

In contrast, APOLLOS utilizes the requirements from Decision Step 1 to find assets and, instead of collecting all alternatives, seeks to exclude any assets that do not fulfill the minimum criteria. Hence, a real-time evaluation based on expert judgement that should result in one to a few asset alternatives for further review. In the case of the safety critical company, the decision-maker browses the available components and exclude the ones that do not fulfill the minimum requirements.

In comparison, ORION Alpha can become time-consuming if many asset alternatives are provided by the involved stakeholders since each asset, as per model description, needs to be fully evaluated based on all metrics. In contrast, APOLLOS allows the decision-maker to make quick judgements if a certain asset may not fulfill some minimum criteria and

disregard any further evaluation. However, this practice is highly susceptible to subjective bias and may therefore lead to the faulty rejection of an asset. ORION Alpha seems to be tolerant to false rejection but it comes at the cost of more time spent.

Decision Step 4: Asset selection - Given the outcome of Decision Step 3, ORION Alpha mandates that a decision is taken based on all collected metrics also taking inputs from all stakeholders identified in Decision Step 1. Once all stakeholders have given their inputs, the decision-maker uses the data to take the most objective decision of what asset to use. This minimizes any biases from the decision and should ensure that the objectively best alternative is chosen according to the perspectives given by all stakeholders. Once a decision has been taken, the decision and the context shall be documented and stored in the decision repository for future use. In the case of the safety critical company, after completing Decision Step 3, the user should have the information required to take an informed and objective decision on what asset and asset origin to choose. In the observed case, the decision was taken to develop an in-house solution to maximize control over the asset due to its tactical importance.

In contrast, APOLLOS implores the decision-maker to take a decision as quickly as possible, disregarding potential bias in such a decision. In the best case, only one asset, or at least asset origin, remains from Decision Step 3 at this step, making the choice easy. If multiple assets are available, the decision-maker decides which asset, according to their best judgement, fits the requirements the best. In the case of the safety critical company, the result, given the requirements from Decision Step 1, should be the same outcome as that for ORION Alpha, i.e., in-house development.

In comparison, since the groundwork for the decision is made in previous steps, both models should be quite equal in time-to-decision at this stage. However, it is still perceived that due to the larger amount of available data available to the decision-maker using ORION Alpha, the decision itself may be more complicated than with the information supplied by APOLLOS.

Decision Step 5: Asset integration - At this step, ORION Alpha assumes that the optimal asset has been identified through the rigorous, and objective, decision process and it can now be implemented. As part of the decision, all properties, contextual models, and other artifacts are saved for use in future decision-making.

In contrast, APOLLOS assumes that the chosen asset may not fulfill all requirements and that adoption/integration of the asset can fail. If such is the case, the asset should fail fast so that the decision-maker can go back to the other alternative assets and use one of them and quickly see if it is more suitable. During this process, if implementation requires further decisions to be taken, APOLLOS allows the decision-maker to do so. This is a feature that ORION Alpha lacks and instead assumes that the final decision will work. In the case of the safety critical company, if an open-source or COTS component was chosen that was later found to not provide enough control, it would, as an example, have been abandoned in favor of in-house development.

In comparison, the overall time to get to integration is perceived faster with APOLLOS. However, this assumes that the decision that was taken is either correct or “good enough.” This could however be false, given a failed decision, implying much additional cost in revising the decision and finding a more suitable asset. Thus, APOLLOS is associated with a cost-effort threat that the model may not be cheaper in the long run despite handling individual asset selection decisions quicker than ORION Alpha.

Summary - In summary, we note that there is a larger upfront evaluation cost associated with ORION Alpha before a decision. This cost is required to ensure objectivity of the decision and that both the needs and assets are analyzed from multiple perspectives. As an effect of this rigorous analysis, the precision of the decision, i.e., the chosen asset's fit for the needs, should be improved. In contrast, APOLLOS quickly provides a decision but with lower precision due to bias, which could result in a failed integration of the chosen asset. Consequently, since ORION Alpha can provide a more precise decision, it is a suitable model for more tactical and strategic decisions, while APOLLOS, due to its possible imprecision, is more suitable for less critical operational decisions. Furthermore, albeit providing support for quicker, individual decisions, in the long-run APOLLOS may be more time-consuming if several consecutive failed decisions are made.

At the Scandinavian safety critical software development company, the observed decision was taken using expert judgement without any decision model or other guidance. The decision, as mentioned, was to develop the image recognition library in-house based on the rationale that it was the only solution that gave the company acceptable control of the asset for long-term maintenance and evolution.

3 Methodology

The first objective of this research is to identify what preference industrial practitioners have regarding decision models for asset selection when given the option to choose between a model with emphasis on high decision precision and a model with emphasis on decision speed. A second objective is to identify, from such preferences, which characteristics of such models are important or influence the practitioners' preferences. These objectives were broken down into the study's research questions, presented in Section 1, which were answered through a survey study that was performed according to the four steps presented in Fig. 3.

At the start of the study, in Step 1, ORION Alpha had already been published (Wohlin et al. 2016) and provided us with a suitable representation of a high-precision decision model. To acquire a model of high speed, we developed the APOLLOS model using the same foundation of knowledge that ORION Alpha was built upon. This foundation includes

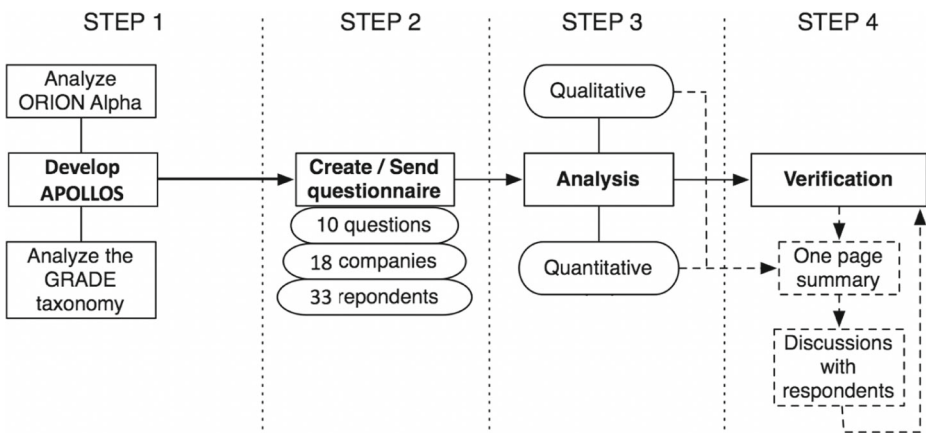


Fig. 3 Visualization of the methodology that used in this work

previous decision-model literature (Berger 2013; Einhorn and Hogarth 1981), and the GRADE taxonomy (Papatheocharous et al. 2015). ORION Alpha itself was also analyzed to get input for APOLLOS. For example, including a knowledge repository in APOLLOS was a result of this analysis. The resulting APOLLOS model was verified within the research team but no external validation, e.g., industrial validation, of the model was done prior to this study.

In Step 2, a questionnaire was constructed that aimed to elicit practitioner preference. The questionnaire was created in Google forms and was divided into two parts. Part 1 described the two models (available as online supplementary material: <http://bth.diva-portal.org/smash/record.jsf?pid=diva2%3A1354531&dswid=-2928>) and part 2 presented the questionnaire questions, described in Table 1. Question Q1 of the questionnaire was binary and aimed at identifying which of the two models was the most similar to the existing

Table 1 The questionnaire questions used in the survey, including their scales. Note that we refer to Decision process 1 (ORION Alpha) and Decision process 2 (APOLLOS) as models in this paper

No.	Question	Scale
Q1	Which of the above described decision-making processes is more similar to the process used for decision-making in your organization?	Binary
Q2	On a scale from 1–10 how useful do you perceive “Decision process 1” to be? (Would it, if adopted, add value to your organisation?)	10-point scale
Q3	On a scale from 1–10 how applicable do you perceive “Decision process 1” to be in your company? (Could it be adopted/used in your organization?)	10-point scale
Q4	On a scale from 1–10 how useful do you perceive “Decision process 2” to be? (Would it, if adopted, add value to your organisation?)	10-point scale
Q5	On a scale from 1–10 how applicable do you perceive “Decision process 2” to be in your company? (Could it be adopted/used in your organization?)	10-point scale
Q6	To what degree do you prefer on of the processes over the other? (The closer to the ends of the scale the more you prefer one process over the other)	10-point scale
Q7	If you got to adopt one of the two processes and personally use it in daily work, which process would you prefer?	Binary
Q8	If you got to choose one of the two processes for your organization to use for decision-making, but not necessarily you personally, which process would you prefer?	Binary
Q9	What challenges/problems do you perceive with “Decision process 1”? (Optional)	Free text
Q10	What challenges/problems do you perceive with “Decision process 2”? (Optional)	Free text

decision-making at the survey respondent's company. The rationale for the question was to acquire insight into the respondents' current work practices, current preferences, and potential bias toward a way-of-working.

Question Q2 aimed to elicit the survey respondents' perception of ORION Alpha's usefulness while Question Q3 aimed to elicit its applicability. These two questions assume that a model can be applicable without being useful but not useful without being applicable. Questions Q4 and Q5 aim to elicit the same information as Q2 and Q3 but for APOLLOS. The rationale behind these four questions were to acquire the level of perceived usefulness of the models. These questions also served as a sanity check of the practitioners' understanding of the questions since it is assumed that the applicability of a model should be rated at least equal, or higher, than the usefulness of a specific model. An assumption that held true in 79% of the responses for ORION Alpha and 72% of the responses for APOLLOS.

Question Q6 aimed to identify the distance in preference between the two models by having the survey respondents' compare them on a single 10-point scale. Hence, if a respondent provided an answer closer to 1, (s)he more strongly preferred ORION Alpha and an answer closer to 10 showed preference toward APOLLOS. This information would further provide a sanity check of the consistency of the practitioners' answers. For example, if a respondent had stated that ORION Alpha was both useful (e.g., answering 8 out of 10 on Q2) and applicable (e.g., answering 9 out of 10 on Q3) but that APOLLOS is neither useful (e.g., 3 out of 10) or applicable (4 out of 10) then it is expected that the respondent would answer Q6 closer to 1 and than 10. Hence, showing preference toward ORION Alpha. Analysis of this assumption showed that it held true in 90% of the responses.

Question Q7 aimed to elicit which model the practitioners would like to use on a daily basis. In turn, question Q8 aimed to elicit which model they thought their company should use. The aim for these two questions was to elicit if a respondent may have found one model applicable and useful but did not want to use it but saw value in their organization using it. These questions thereby gave further insight into the respondents' preferences and if they answered questions Q2, Q3, Q4, and Q5 based on their own subjective view or if they considered the models' usefulness and applicability in a wider, e.g., organizational, perspective. This reasoning was tested on the assumption that if a respondent favored a particular process in Q6, they would like to personally use it as well as want their organization to use it. Analysis of this assumption showed that it held true in 96% of the responses.

Finally, questions Q9 and Q10 were open-ended questions about the respondents' perceived challenges/problems with each model respectively. The rationale for these questions was to get insights into why the respondents had answered the other questions the way they did, e.g., explain their preferences toward one model or the other. Additionally, the qualitative data would serve to identify what characteristics of the models the practitioners focused upon during the comparison. For example, if a respondent perceived a model to be cumbersome by some characteristic (e.g., complexity), it is assumed that they would describe this as a challenge.

Details about the survey are presented in Table 2 using the questionnaire survey guidelines defined by Stavru (2014). As can be seen in the table, the target audience were practitioners with different roles from different domains but that regularly take operational, tactical or strategic software engineering decisions. To improve the response rate, the sampling frame (Stavru 2014) was chosen through one-stage sampling (Fowler 2013) from a set of over 300 personal contacts from the authors' contact network. We opted for personal contacts instead of a larger anonymous group since it allowed us to send personal messages to motivate the respondents to answer the survey. The sample size was chosen through random sampling from the sampling frame. Because random sampling was used, the sample may

Table 2 Characterization of target audience, survey population, and sampling

Characteristic	Description
Target audience and population	Industrial practitioners who participate and experience decision-making and decision-making models in practice.
Sampling frame	Over 300 personal contacts
Sampling method	Randomly from the sampling frame, which was chosen by convenience
Sample size	52 + 10 = 62
Respondents that replied but declined to participate	3
Broken e-mail addresses	3
Response rate	55.9 % ((33 completed + 3 incomplete responses) / 59 delivered e-mails)
Verification survey sample size	10
Response rate	60 %

be subject to sampling error (King and He 2005), e.g., be skewed in the number of respondents in one company compared with another. However, since the sample consisted of 62 individuals from 18 companies, it is perceived to provide an acceptable degree of external validity given the sampling frame (Fowler 2013; King and He 2005). The external validity was also strengthened by the agnostic approach to how recipients of the survey were chosen from different companies, domains, business types, and company sizes and role the recipient had at the company. The questionnaire was sent out in two iterations, first to 52 individuals and later to 10 more from which the final set of 33 responses were collected. This number of responses is considered small but sufficient for the macro-level analysis of what characteristics the practitioners consider in the models. The division in the sample is also well distributed among different roles, which supports the external validity of the results. However, it does present a threat in the sense that each group of roles is much smaller than 33. This threat is considered smaller since the considered characteristics are agnostic of role, indicated by the low variance of answers in the survey. A summary of the variety of the sample can be seen in Table 3, including the number of answers from each company, the companies' domains, and what roles that took part from each company. To preserve the respondents' anonymity, only companies with more than one respondent have been named in the table.

Out of the 62 sent e-mails, 59 were delivered since 3 e-mail addresses were broken. The final response sample from the survey was 33 complete and three incomplete (three participants refused to answer), leaving a response rate of (36/59) 61%. Hence, a response rate over the rule of thumb of 60% and above the average reported response rate of 52.7% in organizational research (Baruch and Holtom 2008).

In Step 3 of the study, as depicted in Fig. 3, the survey results were analyzed to identify the practitioners' preferences. First, a manual sanity check, as discussed previously in Section 3, was performed on the survey responses to see if respondents had answered the questions in a consistent manner. Table 3 shows, in the last column, a consistency metric between the similarity of answers from a specific company. The values, low, medium, and high, were calculated based on the variance between the maximum and minimum sums of all answers from all practitioners of a given company compared with the standard deviation

Table 3 A summary of the number of responses from each company, the development domain, business type, similarity between respondent answers, and roles of respondents

Company	Ans	Domain	Business type	Size	Roles	Similarity of ans.
Axis	5	Camera surveillance systems	Embedded software, Web applications	L	Developers, testers, managers	Medium
Saab	4	Air-traffic management systems	Safety critical SW	M	Developers, managers	Medium
Grundfos	2	Fluid-pump systems	Embedded software	S	Managers	High
ABB	2	Automation control systems	Embedded software	M	Developers, testers	Low
TetraPak	5	Processing plant systems	Embedded software	M	Developers, testers, managers	Medium
Inceptive	4	Software Eng. Consultancy	Consultancy, Mobile/Web/desktop applications	S	Developers, managers	Low
Other (<2 responses / company)	9	Games, Apps, advanced algorithms	Bespoke software, Embedded software, Computer games, Mobile/Web/desktop applications	M	Developers, testers, managers	Low*
(Partially) Unknown	2	Leisure Applications	App. developers	S	Developers, managers	Medium
Summary	33 total		Wide range (No strictly life critical)		Developers, testers, managers	Low–medium

Companies with only one respondent have not been listed to preserve anonymity of the respondents. Size is an estimate of the number of employees at the units of the corporation or company that was surveyed where *S* is less than 50 people, *M* is 50 to 100 people, and *L* is more than 100 people. *Value was expected to be low since this set of responses comes from different domains and companies. *Ans.*, answer; *Similarity of ans.*, quantification of results as described in Section 3

of the sum of all answers (4.514). Hence, by summing up all the answers to all Likert-scale questions for each individual of each company gives a total value of their response, i.e., the sum of questionnaire questions q1 to q8. For example, for Axis, this resulted in an array of five answers from the individual respondents with the values 40, 38, 38, 36, and 44, which implies that the maximum minus the minimum value is 8 (44 – 36). Comparing this value to the standard deviation of all sums of values gives us a measure on the respondents' variation to the norm, i.e., in the case of Axis the value of 8 is compared with the standard variation of 4.514. Thus, if the similarity of answers is denoted high in Table 3, this means that the distance between the minimum and maximum answers was less than one standard deviation apart ($x < 4.514$), medium that the answers were more than one but less than two standard

deviations apart ($4.514 < x < 9.818$), and low that the answers were more than two standard deviations apart ($x > 14.727$). As per our example with Axis, implying that their similarity was “medium” as $4.514 < 8 < 9.818$.

The results were also analyzed using descriptive statistics, where the survey results were plotted to show trends in the data to allow for further conclusions to be drawn. Answers to questions Q2 to Q5 regarding the perceived usefulness and applicability were rendered based on the medium value of acquired responses in spider charts and show the variation between practitioners’ perceptions in different companies. These questions were also analyzed using correlation to investigate if there were any relation between how respondents ranked the models’ usefulness and applicability. First, the results were plotted using a scatterplot, followed by analysis with formal statistics using non-parametric Spearman correlation; results are shown in Section 4.1.

The descriptive statistics indicated that there was a significant difference in terms of usefulness and applicability of the two models. To verify this observation, a non-parametric two-sample Wilcoxon test was also applied at a 95% confidence interval for both usability and applicability. Finally, the answers to Q6 were analyzed using descriptive statistics, plotting the results in bar graphs to visualize how practitioners rank the models compared with one another.

The qualitative data analysis was performed with the free text questionnaire results from questions Q9 and Q10 by using open coding. More specifically, this meant that statements concerning similar drawbacks with a model, e.g., “Too slow” and “Many steps,” were labelled with the code “Speed,” meaning the time required from start to finish to make a decision with the model. In total, 55 statements were collected with an average of 20 words per statement. These statements were clustered into seven keywords that were determined from the data and represent characteristics that practitioners considered when evaluating a decision model; results are shown in Figs. 8 and 9 in Section 5. To validate the soundness of the keywords, they were internally discussed and evaluated by the authors until an agreement of their suitability as a contribution was found. The raw data used in the semantic clustering can be found in the online supplementary material and the resulting keywords can be found in Table 5.

Finally, in Step 4 of the study, the results of the qualitative and quantitative analysis were sent out to all respondents as a one-page report for verification. The respondents were asked to review the results and give additional inputs if they had any. Informal discussions were also held with several of the respondents to get further insights about their view of the two models. These discussions were held open and only served to confirm the identified result of the survey.

4 Results

The results of the study have been divided into two parts and will be presented according to the study’s two research questions.

4.1 RQ1: ORION Alpha compared with APOLLOS

The model preference was evaluated with complementary survey questions to determine to what degree and in what context the participants preferred one model over the other. First, to establish a baseline on how practitioners work with decision-making today, the respondents were asked to compare their current ways-of-working with the processes depicted by

the two models. Results showed that 9 respondents used a decision approach more similar to ORION Alpha than APOLLOS. Inspection of the results showed that most companies of the sample were represented by these 9 respondents, i.e., Axis, Saab, Inceptiv, TetraPak, and several of the anonymous companies. Closer inspection also revealed that the majority of these respondents were in management positions at their respective companies. The remaining 24 respondents used approaches more similar to APOLLOS and, not surprisingly, the majority of these respondents hold developer, tester, and other operational roles.

The models' perceived usefulness was then evaluated in question 2, depicted in Fig. 4 in a radar chart plotted using the median value of the responses of each company. The chart shows that APOLLOS (represented with a solid black line) was generally considered more useful by all respondents and all companies. Inspection of the figure shows that Axis was the most positive to APOLLOS (median 9 out of 10) while Saab was the least positive (median 6.5 out of 10). For context, Axis is an embedded surveillance camera developer, while Saab develops safety critical air management software. Inspecting the figure for perceived usefulness of ORION Alpha shows that ABB was the most positive to the model (median 4 out of 10) and Inceptiv was the least positive (median 2.5 out of 10). For context, ABB develops automation control systems, including safety critical software, while Inceptiv is a smaller software engineering consultancy company.

Closer inspection of the results from individual respondents showed that only two respondents considered ORION Alpha to be more useful than APOLLOS (ORION Alpha ranked 5 and 8 out of 10 and APOLLOS 3 and 6 out of 10). A third respondent perceived

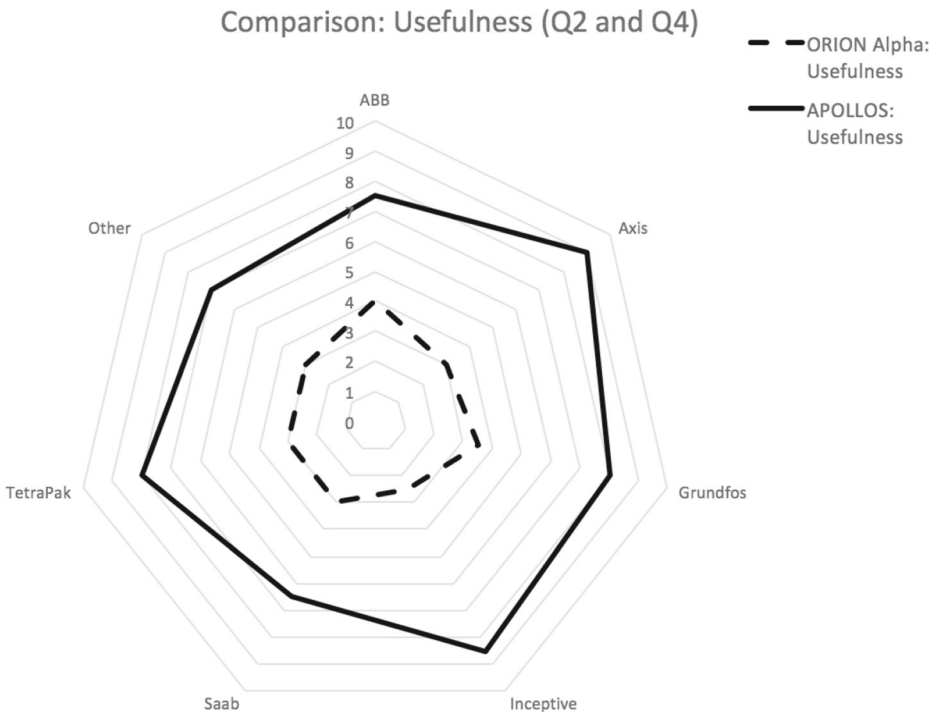


Fig. 4 Radar chart depicting the respondents' answers regarding the perceived *usefulness* of the ORION Alpha (dashed line) and APOLLOS models (solid line). Perception is measured on a 10-point Likert scale from "Not useful" to "Very useful" and has been divided according to the median responses of each company

both models as equally useful (both ranked 8 out of 10). Analysis of who these three respondents showed that all were in managerial positions in their respective companies where they are responsible for both tactical and strategic decisions.

Next, looking at the applicability, depicted once more as a radar chart in Fig. 5, we see more variation in the results. APOLLOS (represented with a solid black line) is still perceived as more applicable in general, but the variance in results between the two models is smaller. For instance, in the case of ABB the median for ORION Alpha’s applicability was 6 out of 10 and for APOLLOS 7 out of 10. Once more, the largest difference was observed from the responses from Inceptive, where ORION Alpha had a median of 3.5 out of 10 and APOLLOS 9 out of 10. Furthermore, in 6 out of the 33 responses, ORION Alpha was considered more applicable than APOLLOS and in one case both models were perceived equally applicable.

Calculating the median of all answers for perceived usefulness and applicability, we see that both the median usefulness and applicability of ORION Alpha were ranked 3 out of 10 by the practitioners overall. For APOLLOS, there was a slight variation, where the usefulness was ranked as 8 out of 10 and the applicability 7 out of 10.

Next, the answers to questions Q2 and Q4 about the usability of the models and questions Q3 and Q5 about the applicability of the models were analyzed using correlation. First, the results were plotted in scatterplots, as shown in Fig. 6a and b. Figure 6a indicates no correlation between how respondents answered questions Q2 and Q4: correlation coefficient $-0.110(p = 0.5415)$. Hence, the results showed no correlation and were also not statistically significant. Figure 6b indicates a negative correlation between the respondents’ answers to questions Q3 and Q5 (correlation coefficient $-0.318(p = 0.071)$), hence

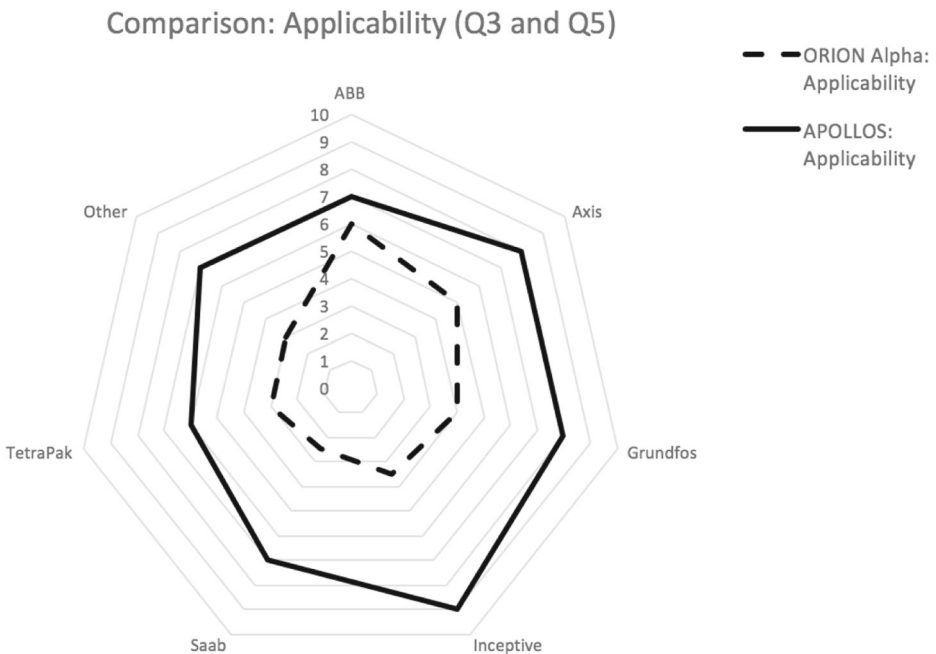


Fig. 5 Radar chart depicting the respondents’ answers regarding the perceived *applicability* of the ORION Alpha (dashed line) and APOLLOS models (solid line). Perception is measured on a 10-point Likert scale from “Not useful” to “Very useful” and has been divided according to the median responses of each company

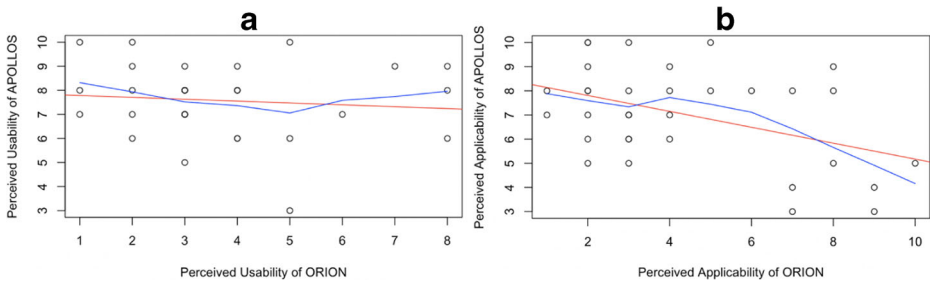


Fig. 6 Scatterplots of the respondents perception of the usability and applicability of ORION Alpha and APOLLOS. The left graph (denoted **(a)**) presents the correlation between the perceived usability of the two models and the right (denoted **(b)**) the perceived applicability of the models. The red line presents the linear regression line of each plot and the blue the locally weighted scatterplot smoothing (Lowess) line, embedded to simplify visual analysis of the plots

a slight negative but not statistically significant correlation. Because of these results, we can draw no conclusion regarding if the respondents answered these questions in relation to one another.

To verify that there is a statistically significant difference between the usability and applicability of ORION Alpha and APOLLOS, two two-sample Wilcoxon tests were performed to test the null hypotheses that there is no difference in the sample means. Hence, if there is any statistically significant difference in the usability of ORION Alpha compared with APOLLOS and there is any statistically significant difference in the applicability of ORION Alpha and APOLLOS. Results for the usability comparison are statistically significant ($p = 3.919e-9$) and we must therefore reject the null hypothesis. Post-hoc power analysis² using the sample mean and standard deviation gives us a power of 1 for our sample size of 33 subjects. For the applicability, the result was once more statistically significant ($p = 6.545e-5$) and we once more reject the null hypothesis. Post hoc power analysis shows that we have a power of 0.997 for our sample of 33 subjects. Hence, there is a statistically significant difference in terms of the practitioners' perceptions on the usability and applicability of the two models.

To strengthen the evaluation of model preference, questionnaire question Q6 asked the practitioners to choose one of the two models in a 10-point scale where 1 indicated strong preference toward ORION Alpha and 10 strong preference toward APOLLOS. The result of the evaluation is shown in Fig. 7, which shows that 11.6% of the respondents favored ORION Alpha and 88.4% favored APOLLOS. The median value of the 3 respondents that favored ORION Alpha was 2 out of 10 while the median for the 30 respondents that favored APOLLOS was 8 out of 10. The total median of all 33 answers was also 8 out of 10. Thus, 30 out of 33 respondents showed preference toward APOLLOS (representative of a quicker but less precise model) compared with 3 respondents that showed preference toward ORION Alpha (representative of a slower but more precise model).

Finally, to complement the results about preference even further, the respondents were asked to choose which model they would prefer to work with on a daily basis and which model they felt was more appropriate for their organization to use. The result reflects the previous results that 30 out of the 33 respondents wanted to use APOLLOS rather than

²We recognize that post hoc power analysis is bad practice but decided to include this evaluation for transparency of the statistical analysis.

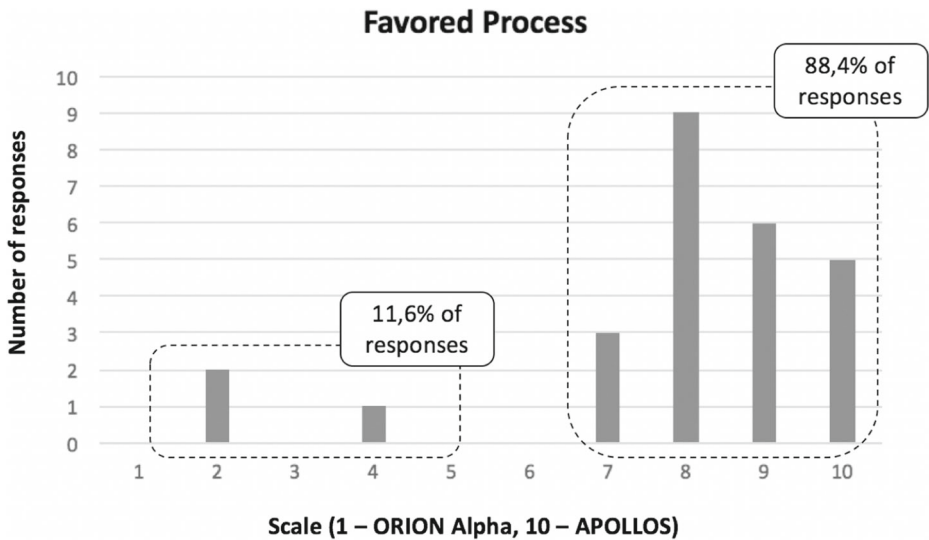


Fig. 7 Bar graph depicting which of the ORION Alpha and APOLLOS the respondents perceivably preferred. An answer closer to 1 on the scale indicates preference toward ORION Alpha and an answer closer to 10 a preference toward APOLLOS

ORION Alpha in their own work. Additionally, 29 out of the 33 respondents wanted their organization to use APOLLOS. The respondent that changed their opinion was a developer of a core IT team at TetraPak, which is an older organization that has a long tradition of developing automated solutions and embedded systems.

From these results, taking into account that APOLLOS is representative of faster but less precise decision models for asset selection and ORION Alpha is representative of slower but more precise decision models, we come to the conclusion that:

1. **C1:** When given the choice between a decision model for asset selection with high precision or high speed, the industrial practitioners will show preference toward the model of high speed.

4.2 RQ2: Perceived drawbacks and challenges

To answer research question two, the free-text answers of the survey, survey questions Q9 and Q10, were analyzed to identify what the main drawbacks of each model are perceived to be. A summary of the key challenges has been presented in Table 4, formulated through synthesis of the acquired statements from the practitioners.

ORION Alpha A recurring statement about ORION Alpha was that it was perceived slow and a total of 11 statements could semantically be linked to the model’s lack of speed. The practitioners’ concern about the speed can be associated to the modern, agile, ways-of-working that require rapid response to change and thereby quick decisions. One comment to support this observation was, “It’s (ORION Alpha) way too slow. The tech we work with moves on very quickly.” However, the precision of the model was not overlooked by the survey respondents: “model 1 (ORION Alpha) appears to be quite time and resource

Table 4 Summary of the key challenges that the practitioners perceive with the two models

Model	Challenge	Description
ORION Alpha	Slow	The model was perceived large, with many time consuming steps and therefore long lead times to decision.
	Complex/ stringent	The model includes a process with many steps that require data acquisition as well as several knowledge databases that are perceived as complex.
	Rigid with many steps (not agile)	The model's process is perceived to have dependencies among the steps that make it rigid and difficult to manage for certain decisions and contexts.
	Inflexible (not suitable for day-to-day decisions)	Because of its size and complexity the model is not perceived suitable for smaller day-to-day decisions because the time to decision would be too long.
APOLLOS	Imprecise	The model is perceived to not provide enough time to evaluate the options, instead quick decisions are premiered, which is perceived as imprecise.
	Wasteful (failure prone)	Because of the model's focus on speed and its impression it is perceived to lead to faulty decisions that may require sunken efforts to be thrown away, which is considered wasteful.
	Too quick	The model is perceived to stress the decisions, not providing enough time to evaluate the options in a rigorous enough way.
	Inflexible (not suitable for larger decisions)	The model is not perceived to scale for larger decisions, especially decisions where the outcome must be correct/precise.
	Does not keep people in the loop	The model is perceived to not support collaboration among multiple stakeholders.

consuming passing through all the steps. Initially it might be hard to get appreciation for the effort required for a single decision, but the result of the actual decision would probably be closer to the target, and the knowledgebase will prove valuable over time to help make the model faster / easier and iteratively improve the accuracy of decisions made.”

Another commonly stated drawback, associated with the first, was that the model felt too academic, complex, and heavy to use: “Too heavy/ academic/ bureaucratic. My staff would probably reject this model after having read only a third of the description text, since it seems overly structured. Might be ok for extremely complex decisions with a lot at stake (deciding if going to build high-speed trains or a new airport), but not applicable for most decisions in daily work.” The complexity of the model was therefore inferred to reduce its use for decisions in an agile context: “Way too complex, reminds me of an old school waterfall project model.” This was further supported by statements like, “Too complex... you would not be able to make any decision at all.” and “High administration cost and non agile.” Hence, we observe that the process was perceived to be too stringent. An observation further supported by the statement: “(ORION Alpha) It's very formal and process bound.”

However, the respondents also realized the need to consider the decision context, including the time to decision and the type of the decision, e.g., operation, tactical, or strategic. For instance, one respondent stated: “Seems like the type of formal decision process (ORION Alpha) used for making big decisions where you only get one chance at succeeding (e.g. government decisions on using nuclear power or similar!),” thus implying the model’s usability for more important, larger decisions, which has the resources for a longer evaluation. However, for smaller decisions, “This process (ORION Alpha) is not very applicable when dealing with smaller everyday decisions due to the complexity and effort, but for “higher” up decisions in the organization it would prove beneficial.” The core reason once more being the model’s speed and complexity for smaller everyday decisions: “Too slow/complex, but could be valid for non-time critical decisions.” A synthesis of these statements indicates that ORION Alpha is more suitable for tactical and strategic decisions but less suited for operational decisions.

APOLLOS The main drawback perceived with APOLLOS was that the model would lead to less precise, and therefore wrong, decisions. For instance the respondents stated that: “not as precise as model 1 (ORION Alpha)” and “Process 2 (APOLLOS) will not give the level of accuracy as Process 1 (ORION Alpha) due to the shortened decision process but will improve the decisions made on a more daily basis.” Some respondents even stated that: “There is a risk for too quick decisions and that you do the wrong thing and lose time.” and “(APOLLOS is) Wasteful (work put in might be thrown away if the tool/(asset)/etc. was not suitable).” The core challenge is thereby that with APOLLOS you need to accept that some decisions may be wrong. However, as stated by one respondent, “You have to be allowed to fail. If you only got one shot at something this might not be your best option.” The speed was also lifted as a concern since it may result in users overlooking existing solutions: “We might miss knowledge that already is available at the company, since the process might be too fast.”

Other drawbacks were connected more to contextual factors such as the type of the decision: “How much knowledge you need to make the decision is based on the complexity and the consequences of the decision. Therefore, the model needs to take into account the decision’s complexity.” While ORION Alpha was associated more with larger decisions, APOLLOS was by the respondents associated with smaller ones: “(APOLLOS) Might be too simple for extremely complex decisions.”

Other contextual factors that were brought up included human factors, for instance that it might be difficult to use APOLLOS in an environment with many people: “(APOLLOS could be) Difficult to use when it involves large changes and a lot of people.” Another respondent also pointed out that models, or processes, do not take the decisions, it is the human(s) in the loop that matter: “The people involved in making the decision is much more important than the process used. You need to make sure the right people (often users - not “stakeholders” which owns the budget) are included and have the decision power.” An observation drawn from these statements is that APOLLOS seems more suitable for operational decisions.

In summary, the results lead us to the following answer to research question 2:

1. **C2:** The respondents associate ORION Alpha with high cost but also recognize the model’s usability for more complex and larger decisions (tactical and strategic decisions) where precision is important. In contrast, APOLLOS is considered to lack precision but because of its speed it is perceived applicable for more day-to-day (operational) decisions.

5 Discussion

In this section, we will discuss the results of the study and the characteristics that we identified that affect the practitioners' preference toward a model.

5.1 Model characteristics

In Section 4.2, we observed that the speed and precision of the decision models are important. These are the key characteristics on which the ORION Alpha and APOLLOS were compared. However, a semantic analysis with open coding of the free-text answers from the respondents identified in total seven characteristics that the practitioners found important, summarized in Table 5. These characteristics were identified based on the stated challenges with the models, identified through clustering of statements as presented in Figs. 8 and 9. Note that the statements have only been associated with the characteristic they are perceived to support the most, i.e., several statements support several of the characteristics. For example, the statement, "Too slow/complex, but could be valid for non-timecritical decisions," could have been associated with both complexity and speed. The clusters have been color-coded depending on for which model the statement was given, i.e., blue lines for ORION Alpha and orange lines for APOLLOS. As can be seen, the different characteristics were associated with one or the other model, where characteristics such as complexity, speed, flexibility, and agility were more prominent for ORION Alpha and characteristics such as people, precision, and failure were associated with APOLLOS. Once more, note that the statements were clustered based on the primary characteristic that they discuss, as presented in Section 3, even though several statements could be mapped to several characteristics.

The continuation of this section will describe the identified characteristics and discuss them from the context of operational, tactical, and strategic decisions (Moe et al. 2012). Definitions of the three types of decisions are presented in Table 6.

Complexity This characteristic was presented from two perspectives by the practitioners. The first regards the complexity of the model both in terms of appearance and of perceived complexity of use due to rigor and number of actual components, e.g., steps or repositories. While the former can be mitigated through presentation, e.g., by changing the level of abstraction of the model, the latter requires more concern and understanding of the user's needs for the model. Industrial practitioners strive for "good enough" solutions because it minimizes cost but what is good enough varies with the decision being made. The concept of good enough is therefore problematic because it infers a level of risk that for minor decisions might be easier to justify than for larger decisions. Decision support models thereby have a threshold to how many correct decisions they make per faulty decision, which can be directly correlated to their rigor, i.e., a more rigorous and objective model should be more often correct than a less rigorous or subjective model.

This leads us to the second perspective of complexity that was mentioned, which refers to the aforementioned context, or level, of the decision. In this work, the decision models were designed for asset selection, a type of decision that is generally quite small and performed on a day-to-day basis (operational) but sometimes has larger architectural impact (tactical) but seldom has any larger impact on the product (strategic) (Moe et al. 2012). As such, asset selection decisions emphasize speedy decisions and, because of their lower impact on production, the precision of the decision is not crucial. This explains why APOLLOS was favored since it is perceived to better fulfill these requirements. A summary of the perceived application contexts of the two models has been presented in Table 6.

Table 5 Summary of the main needs and derived required characteristics for these needs that were identified through qualitative analysis of the survey results

No.	Characteristic	Industrial need	Characteristic description	Observed in model
1	Complexity	The decision model must be easy to understand	A decision model shall be managed in terms of presentation and included components for it to easily be digested and understood	OA, A
2	Speed	The execution time of a decision model must be short	A decision model shall allow the user to quickly acquire a “Good Enough” decision	OA, A
3	Flexibility	The decision model shall be applicable for both small and large decisions	A decision model must be flexible enough to handle both smaller and quick decisions as well as larger and complex decisions	OA, A
4	Agility	The decision model shall support the user traversing its steps in any order	A decision model must not enforce how its model shall be traversed but rather allow the uesr to traverse the model in a way that best suits the decision-making context	OA, A
5	People	The decision model shall support the people	A decision model is used by people who are the ones that take the final decision, the model shall be designed to support them.	OA, A
6	Precision	The decision model shall give precise support and a well-defined outcome	A decision model must be strict enough to give a well-defined, precise output of which decision to make	A
7	Failure	The decision model shall support failure	A decision model shall provide a decision that to some degree works, at least for some time, but decisions sometimes fail and that must be dealt with quickly (fail fast)	A

Supporting quotes was calculated as the number of quotes coded with each characteristic. OA, ORION Alpha; A, APOLLOS

Speed This characteristic is important to consider when creating/adopting/using a decision model, especially in unison with good enough decisions since it affects the decision’s cost. In short, speedier and correct decisions are associated with lower cost but speedy and incorrect decisions may greatly increase cost. Hence, speedier decisions are associated with a higher risk and, as discussed in relation to complexity, the threshold for a number of correct versus incorrect decisions are then also correlated with the speed. The reason is because faster decisions have less data to base the decisions on and therefore have a larger statistical



Fig. 8 Semantic clustering of statements from the open questions. Blue lines refer to statements about ORION Alpha and orange lines statements about APOLLOS

chance of being wrong. As a consequence, the contextual factors of the decision once more come into play where speedy but potentially faulty decisions may be acceptable for operational decisions but not acceptable for strategic decisions. Thus, for any specific decision model, it is important to consider what type of decision it supports, how critical the decision is, and what resources could perceivably be devoted to said decision. In the case of asset selection, being a mostly operational decision, speed is important and faulty decisions tolerated. However, for more tactical or strategic decisions, the precision of the decision becomes more impactful since a faulty decision will be associated with a larger cost. This line of reasoning also implies that speed becomes less important for higher level decisions while precision grows in importance. It is possible that other characteristics are equally important depending on the contextual level of the decision, but since this study has focused on speed and precision such discussion is left to future work.

Flexibility This characteristic is heavily influenced by the decision levels since higher flexibility implies that the model is usable for more types of operational, tactical, and strategic decisions. However, to achieve generalizability, the model also has to become more abstract, which in turn makes the model less supportive to the decision-maker and could detrimentally affect both speed and decision precision. Other factors could be affected as well but in this work we focus on speed and precision. Consequently, there is a trade-off to be made when developing, adopting, or using a decision model that regards model precision compared with its applicability for different levels of decisions. This trade-off should be balanced based on the return on investment of the model's use and consider (1) the cost of using the model, (2) the level of the decision, and (3) the cost of a faulty decision, i.e., the risks taken when making a good enough decision. If the cost of using the model for a

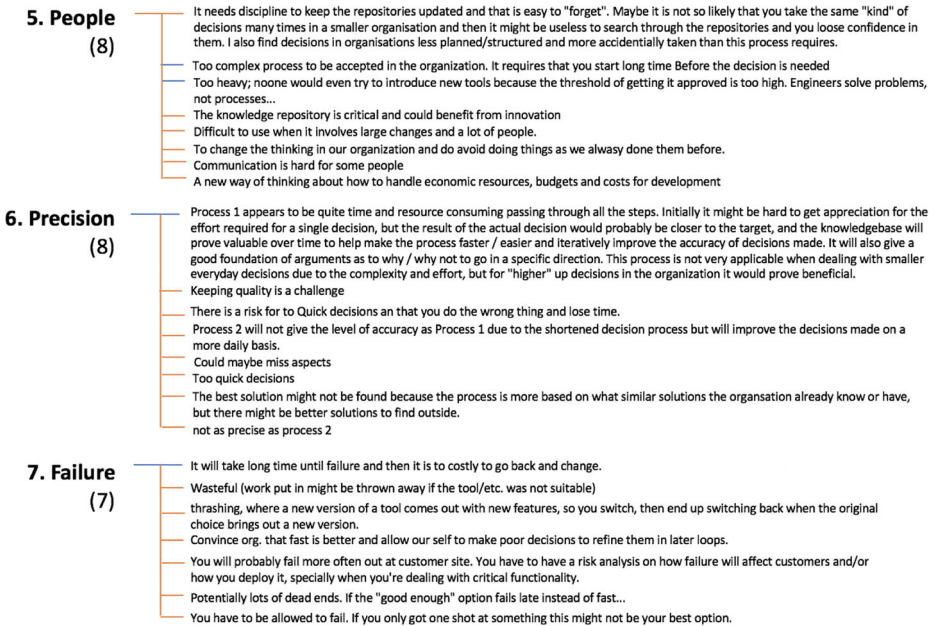


Fig. 9 Semantic clustering of statements from the open questions. Blue lines refer to statements about ORION Alpha and orange lines statements about APOLLOS

decision on a given level outweighs the cost of failing to take the right decision, the model should not be used on that level. It is important to remember that decision models are meant to provide decision support, not to take the decision, and any unnecessary time collecting information that is not required is thereby pure waste. However, the same can be said if the model does not provide enough information for a correct, or at least good enough, decision to be taken. As such, for any good enough decision to be taken, an estimation or measure of the risk of failure should be considered.

Agility While flexibility refers to the type of decision the model supports, the agility of the model refers to the process of how said decision is taken. This characteristic is important to consider because, even though it affects the flexibility, it also affects how decisions are taken, when and by who. In short, a flexible model is perceived to be adaptable to a certain decision context where more/less stakeholders, information, time, or other resources may be available. This implies that more responsibility for the decision's precision is

Table 6 Types of decision levels the different models are perceived more or less suitable for

	Decision model	Decision level description
Operational	APOLLOS	Decisions that are performed every day and require little effort with smaller impact.
Tactical	ORION Alpha	Decisions that are performed periodically and require some effort but range from small to high impact.
Strategic	ORION Alpha	Decisions that are performed less frequently but require considerable effort due to their high impact.

transferred to the decision-maker or stakeholders that provide the decision support information. Meanwhile, this also means that the model can be tailored to support more efficient decision-making given a certain context or decision. Once more, the developer, adopter, or user of the model needs to consider how this characteristic is represented in the model and how it affects the model's usability and cost.

People Although decision models provide support, decisions are ultimately taken by individuals. Developers, adopters, and users of decision models therefore need to consider which stakeholders are part of the decision and either tailor the model to these stakeholders, or choose an appropriate model, accordingly. This aspect is also emphasized in the GRADE taxonomy (Papatheocharous et al. 2015). Furthermore, this characteristic implies that decision model developers must elicit the needs of the decision-makers to make the models usable for the users' needs. This lack of elicitation could explain ORION Alpha's lack of practitioner preference since it seems that the model's rigor and size make it misalign with the decision level where it should be used.

Precision This characteristic should always be optimized since faulty decisions can be associated with considerable cost. However, this cost is also associated with the decision level where operational decisions may have minor to no impact while a strategic decision can have devastating impact. Hence, although the support for a decision, and the decision itself, should be good enough, good enough still implies that the correct decision was taken given the applied effort. As a consequence, a model that provides quick but continuously faulty decisions is still less useful than a model that requires marginally more than good enough effort but correct decisions. Once more, it comes down to the return on investment of using the model contra the cost of taking a faulty decision, i.e., the risk taken to save time or money. Similarly, this characteristic places requirements on the model developer, adopter, or user to consider the contextual factors and evaluate the return on investment of having higher decision precision.

Failure Good enough, per definition, is not perfect and thereby carries a natural risk of failure, i.e., leading to faulty decisions. The chance of failure can be mitigated with a more stringent process, as in ORION Alpha, but that also mitigates the model's perceived applicability for smaller, e.g., operational, decisions. Alternatively, the model can accept that some decisions will be faulty and propose means to deal with them, as in APOLLOS. Regardless, a risk-aware mindset should be employed when developing/using the models to gauge what is acceptable risk for a given type of decision. Once more, the context and decision level becomes important for the developer, adopter, or user of the model to estimate what the acceptance of a failure may be for a certain decision. As a guideline, for operational decisions, the failure acceptance is higher and for strategic decisions the failure acceptance is lower, as summarized in Table 6.

5.2 Threats to validity

This section will discuss the threats to the validity of this research following the guidelines set by Runeson and Höst (2009).

Internal validity regards the appropriateness of the casual relationships that can be observed from the results and thereby the study results' ability to support its conclusions. Although the study could clearly show that the practitioners favored the APOLLOS model over ORION Alpha, there are a few threats to consider.

First, the questionnaire survey only measured the practitioners' perceptions of the models, not if the models were actually applicable or useful. This implies that other factors, like their bias toward smaller models, could be the casual cause of the results rather than the conclusion we have drawn that APOLLOS is a more favored model for its characteristics. However, we deem this threat to be smaller due to how many respondents favored APOLLOS and the practitioner statements in the open questions.

Second, although a questionnaire survey design is suitable for collecting data from a diverse sample, it limits researcher control, e.g., participants cannot ask for questions to be clarified. As such, it is possible that some of the questionnaire questions, or model descriptions, were misinterpreted, thus leading to faulty answers that can introduce flaws in the data. To mitigate this threat, the questions were designed to give multiple perspectives of the participants' perceptions to enable triangulation. Results of the study were also verified by sending the results to the participants of the survey, resulting in no feedback, implying that the results of our work is correct.

Third, the characteristics were gathered from the answers of the open-ended questions of the survey. These answers may have been misinterpreted and/or are subject to researcher bias. To mitigate this threat, only characteristics that had strong support through multiple comments were stated in the results. However, this threat could have been avoided if the questionnaire guide had explicitly asked the respondents to give their perception of the models' characteristics. The choice to omit this explicit question was however intentional not to steer the respondents and allow a broader spectrum of challenges to be elicited.

Fourth, the sample size ($N_{\text{survey1}}=33$) is small and split across many companies and domains. This may affect the level of granularity at which the results can be presented, e.g., in the discussion of operational, tactical, and strategic decisions, that have been linked to different roles in the sample. However, as the results are highly consistent across all roles, we consider this threat to be minor and perceive the results to sufficiently support our conclusions, particularly since the study's main contribution, the seven characteristics, are quite general albeit important to be aware of in this type of research.

Finally, there is a threat of potential sample bias in terms of the respondents' roles and thereby the types of decisions that the individuals are associated with. However, the sample included roles ranging from developers, managers, and even company CEOs, which should have provided a broad enough spectrum of responses. Regardless, this factor cannot be overlooked and future research should therefore aim to verify the results and also verify the mapping between characteristics and decision levels.

External validity refers to the generalizability of the results that can in this study be questioned due to the small number of survey responses ($N_{\text{survey1}}=33$). However, as the study sought a domain agnostic result, the measure of greater importance for generalizability is the number of responding companies ($N_{\text{companies}}=18$). This high number, joined with the breadth of roles participating in the survey and consistent results on preference, applicability, and usability, in both survey instances, implies that the results are externally valid. However, the results are still based on perception and it is therefore unclear if the same results would be acquired after actual use of the models. In addition, the resulting characteristics are affected by the same limitation and should therefore be further empirically evaluated to improve their validity. Regardless, the reported characteristics, and their connection to the decision levels, are logically sound and thereby pass initial screening. Finally, this work focuses on asset selection for software development and, although the results may be applicable for other types of decision-making, this work makes no such claims. We do however encourage research to validate our results for other types of decision models in software engineering or other domains.

Construct validity refers to the research context and the study design's ability to capture results that in a valid way can answer the study's research questions. The objective of the study was to capture the practitioners' perceptions of the two decision models and from these perceptions identify characteristics of the models the practitioners found important. We believe the study achieves this objective and provides us with a broad set of data from practitioners from many different companies, roles, and types of software systems. We also sent out a one-page summary of the results to the practitioners for verification, and received no feedback on the correctness of our analysis. We therefore assume our results to be valid for the context and companies that took part of the study.

Reliability refers to the ability to replicate or determine the validity of the results based on the study's description. To facilitate such evaluation, the paper has provided detailed descriptions about the questionnaire and the sample group. The raw, anonymized data collected and used in this study have also been made available as online supplementary material: <http://bth.diva-portal.org/smash/record.jsf?pid=diva2%3A1354531&dswid=-2928>. This allows any researcher to see the raw data and verify, or replicate if needed, the presented results.

6 Conclusions

Deciding what technical asset to select when provided with many asset alternatives and the options to develop, outsource, use open-source, or buy COTS assets can affect a product's competitive advantage in the market. Decision support models, such as ORION Alpha and APOLLS, aim to help practitioners in these decisions. However, the models' characteristics affect their fit for different decision levels and thereby their usability and industrial acceptance.

In this study, we have performed an industrial questionnaire survey where industrial practitioners' preference of the ORION Alpha model, which emphasizes decision precision, is compared with APOLLOS, which emphasizes speedy decisions for asset selection. The results clearly show that practitioners prefer APOLLOS which can be explained by the model's better fit for asset selection as an operational decision. However, APOLLOS' lack of precision is presented as a challenge for more critical decisions (tactical and strategic decisions) and support is given that ORION Alpha is a better fit in these circumstances. This conclusion provides support to the established conclusion that the decision context must be considered for a decision model's applicability (Einhorn and Hogarth 1981), thus implying that none of the models is optimal for all types of decisions.

Deeper analysis of the survey's free-text responses enables us to identify seven characteristics that the practitioners consider in their comparison. These characteristics represent the study's main contribution and are model complexity, speed, flexibility, agility, people, precision, and failure, which for varying decision levels need to be considered in different ways. These characteristics complement existing work (Mistree and Allen 1997) and can be used for both specification and classification of decision models. For example, for operational decisions, a failed decision is perceived to have less impact, i.e., risk, and characteristics such as improved speed at the cost of lower decision precision are therefore less of a challenge. However, for strategic decisions, decision precision is crucial because failure will have a significant impact, implying that speed must be sacrificed and complexity increased. Consequently, decision model creation requires trade-offs of different characteristics that should be determined based on the intended use of the decision model and its stakeholders' needs.

Concretely, a decision model aimed at supporting operational decisions should have low complexity, high speed, high flexibility, high agility, few people involved, lower but good enough precision, and capability to deal with possibly faulty decisions. In contrast, a strategic decision model can have high complexity, low speed, lower flexibility, lower agility, many people involved, and high precision but should ensure that faulty decisions are not taken. For tactical decisions, a middle ground of these examples is perceived the most suitable. However, contextual factors also play a role, meaning that these statements are guidelines for what practitioners expect. Regardless, the choices made when designing a decision model will influence the model's industrial acceptance, where a model that requires more than good enough effort for a given decision type or level will be rejected by the practitioner. The target stakeholder of the decision model must therefore be considered since the results of this study indicate that practitioners in managerial roles showed a slighter larger preference toward ORION Alpha, i.e., a more stringent model, than other stakeholders.

Finally, decision models only provide support to the decision-maker who will take the decision in the end, but most decisions are still based on decision-makers own expert judgment (Chen et al. 2008). The decision model may provide objective inputs for this decision and clearly show which asset alternative, according to any and all metrics, is the most suitable. However, the decision-maker may choose to ignore this information and base the decision on other, more subjective, factors. As researchers, we like to propose solutions that use purely objective metrics derived from quality, time, or perhaps cost that lead to the best decision in a rational way. In addition, researchers tend to neglect other practical considerations such as time and resource constraints and assume that practitioners will spend the time required to make the objectively best decision. However, in practice, politics and human factors, such as fear, laziness, and self-perseverance (Borg et al. 2018; John et al. 2005; Guo et al. 2011), also have to be taken into account since they lead to irrational behavior. A potential direction for decision model research should therefore be to look further into the core human factors that affect the decision-maker and their final asset selection. Such research is also warranted by related software engineering research into, for instance, how human factors affect cost estimation (Magazinius et al. 2012), process improvement (Lenberg et al. 2017), and behavioral software engineering (Lenberg et al. 2015).

In summary, this study provides a contribution in filling the gaps in decision models for them to be accepted in practice. Our results thereby provide guidance for future development of decision models for asset selection that enable researchers to improve the models' fit for purpose and industrial acceptance.

Acknowledgments We would like to thank our colleagues in the ORION project for fruitful discussions and the external reviewers that have helped improve the paper.

Funding information Open access funding provided by Blekinge Institute of Technology. This work is supported by a research grant for the ORION project (reference number 20140218) from The Knowledge Foundation in Sweden.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- van der Linden, F., Lundell, B., Marttiin, P. (2009). Commodification of industrial software: A case for open source. *IEEE software*, 26(4).
- Badampudi, D., Wohlin, C., Petersen, K. (2016). Software component decision-making: In-house, oss, cots or outsourcing—a systematic literature review. *Journal of Systems and Software*, 121, 105.
- Ssaed, A., Kadir, W., Hashim, S.Z.M. (2012). Metaheuristic search approach based on in-house/out-sourced strategy to solve redundancy allocation problem in component-based software systems. *International Journal of Software Engineering and Its Applications*, 6(4), 143.
- Vale, T., Crnkovic, I., de Almeida, E.S., Neto, P.A.d.M.S., Cavalcanti, Y.C., de Lemos Meira, S.R. (2016). Twenty-eight years of component-based software engineering. *Journal of Systems and Software*, 111, 128.
- Huhns, M.N., & Singh, M.P. (2005). Service-oriented computing: Key concepts and principles. *IEEE Internet computing*, 9(1), 75.
- Breivold, H.P., & Larsson, M. in Software Engineering and Advanced Applications, 2007. 33rd EUROMICRO Conference on (IEEE, 2007), pp. 13–20.
- Moe, N.B., Aurum, A., Dybå, T. (2012). Challenges of shared decision-making: A multiple case study of agile software development. *Information and Software Technology*, 54(8), 853.
- Petersen, K., Badampudi, D., Shah, S.M.A., Wnuk, K., Gorschek, T., Papatheocharous, E., Axelsson, J., Sentilles, S., Crnkovic, I., Cicchetti, A. (2018). Choosing component origins for software intensive systems: In-house, cots, oss or outsourcing?—a case survey. *IEEE Transactions on Software Engineering*, 44(3), 237.
- Li, J., Bjørnson, F.O., Conradi, R., Kampenes, V.B. (2006). An empirical study of variations in cots-based software development processes in the norwegian it industry. *Empirical Software Engineering*, 11(3), 433.
- Cortellesa, V., Marinelli, F., Potena, P. (2006). in European Workshop on Software Architecture (Springer), pp. 66–81.
- Potena, P. (2007). In Proceedings of the the 6th joint meeting of the European software engineering conference and the ACM SIGSOFT symposium on The foundations of software engineering (ACM), pp. 583–586.
- Stewart, K.J., Ammeter, A.P., Maruping, L.M. (2005). In Proceedings of the 38th annual hawaii international conference on system sciences (IEEE), pp. 197c–197c.
- Rudzki, J., Kiviluoma, K., Poikonen, T., Hammouda, I. (2009). In 2009 35th Euromicro conference on software engineering and advanced applications (IEEE), pp. 11–19.
- Petersen, K., Badampudi, D., Shah, S.M.A., Wnuk, K., Gorschek, T., Papatheocharous, E., Axelsson, J., Sentilles, S., Crnkovic, I., Cicchetti, A. (2017). Choosing component origins for software intensive systems: In-house, cots, oss or outsourcing?—a case survey. *IEEE Transactions on Software Engineering*, 44(3), 237.
- Starmer, C. (2000). Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. *Journal of economic literature*, 38(2), 332.
- Wohlin, C., Wnuk, K., Smite, D., Franke, U., Badampudi, D., Cicchetti, A. (2016). In International conference of software business (Springer), pp. 1–15.
- Badampudi, D., Wnuk, K., Wohlin, C., Franke, U., Smite, D., Cicchetti, A. (2018). A decision-making process-line for selection of software asset origins and components. *Journal of Systems and Software*, 135, 88.
- Martens, B., & Teuteberg, F. (2012). Decision-making in cloud computing environments: a cost and risk based approach. *Information Systems Frontiers*, 14(4), 871.
- Papatheocharous, E., Petersen, K., Cicchetti, A., Sentilles, S., Shah, S.M.A., Gorschek, T. (2015). In Proceedings of the 2015 European conference on software architecture workshops (ACM), pp. 48.
- Dave, M.S., Patel, M.H., Prajapati, M.A. (2018). Non-functional testing in agile development. *IJIRT* 5(7).
- Garg, R. (2017). A systematic review of cots evaluation and selection approaches. *Accounting*, 3(4), 227.
- Maiden, N.A., & Ncube, C. (1998). Acquiring cots software selection requirements. *IEEE software*, 15(2), 46.
- Ven, K., & Verelst, J. (2006). In IFIP international conference on open source systems (Springer), pp. 111–122.
- Li, X., Li, Y.F., Xie, M., Ng, S.H. (2011). Reliability analysis and optimal version-updating for open source software. *Information and Software Technology*, 53(9), 929.
- Yang, C., & Huang, J.B. (2000). A decision model for is outsourcing. *International Journal of Information Management*, 20(3), 225.

- Borg, M., Chatzipetrou, P., Wnuk, K., Alégroth, E., Gorschek, T., Papatheocharous, E., Shah, S., Axelsson, J. (2018). Selecting software component sourcing options: Detailed survey description and analysis.
- Strigini, L. (1996). Limiting the dangers of intuitive decision making. *IEEE Software*, 13(1), 101.
- Aurum, A., & Wohlin, C. (2003). The fundamental nature of requirements engineering activities as a decision-making process. *Information and Software Technology*, 45(14), 945.
- Berger, J.O. (2013). Statistical decision theory and Bayesian analysis (Springer Science & Business Media).
- Einhorn, H.J., & Hogarth, R.M. (1981). Behavioral decision theory: Processes of judgement and choice. *Annual review of psychology*, 32(1), 53.
- Ruhe, G. (2002). In International workshop on learning software organizations (Springer), pp. 104–113.
- Cusumano, M.A. (2004). The business of software: What every manager, programmer and entrepreneur must know to thrive and survive in good times and bad (Simon and Schuster).
- Franke, U., & Buschle, M. (2016). Experimental evidence on decision-making in availability service level agreements. *IEEE Transactions on Network and Service Management*, 13(1), 58.
- Boehm, B.W., Madachy, R., Steece, B., et al. (2000). Software cost estimation with Cocomo II with Cdrom (Prentice Hall PTR).
- Stavru, S. (2014). A critical examination of recent industrial surveys on agile method usage. *Journal of Systems and Software*, 94, 87.
- Fowler, F.J. Jr. (2013). Survey research methods (Sage publications).
- King, W.R., & He, J. (2005). External validity in is survey research. *Communications of the Association for Information Systems*, 16(1), 45.
- Baruch, Y., & Holtom, B.C. (2008). Survey response rate levels and trends in organizational research. *Human relations*, 61(8), 1139.
- Runeson, P., & Höst, M. (2009). Guidelines for conducting and reporting case study research in software engineering. *Empirical software engineering*, 14(2), 131.
- Mistree, F., & Allen, J.K. (1997). Position paper optimization in decision-based design, Optimization in Industry, Palm Coast, FL, Mar pp. 23–27.
- Chen, W., Li, J., Ma, J., Conradi, R., Ji, J., Liu, C. (2008). An empirical study on software development with open source components in the chinese software industry. *Software Process: Improvement and Practice*, 13(1), 89.
- John, M., Maurer, F., Tessem, B. (2005). Human and social factors of software engineering: workshop summary. *ACM SIGSOFT Software Engineering Notes*, 30(4), 1.
- Guo, P.J., Zimmermann, T., Nagappan, N., Murphy, B. (2011). In Proceedings of the ACM 2011 conference on Computer supported cooperative work (ACM), pp. 395–404.
- Magazinius, A., Börjesson, S., Feldt, R. (2012). Investigating intentional distortions in software cost estimation—an exploratory study. *Journal of Systems and Software*, 85(8), 1770.
- Lenberg, P., Tengberg, L.G.W., Feldt, R. (2017). An initial analysis of software engineers' attitudes towards organizational change. *Empirical Software Engineering*, 22(4), 2179.
- Lenberg, P., Feldt, R., Wallgren, L.G. (2015). Behavioral software engineering: a definition and systematic literature review. *Journal of Systems and software*, 107, 15.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Emil Alégroth is a software engineering researcher at Blekinge Institute of Technology and affiliated with Chalmers University of Technology. He has for several years worked with industrially close research while operating as the Chief Executive Officer (CEO) of a spin-off company from his PhD studies. His research interests include decision-making, human factors, and in particular automated software testing.



Tony Gorschek is a professor of software engineering at Blekinge Institute of Technology and parttime at Chalmers. He has over 10 years industrial experience not only as a CTO, senior executive consultant, and engineer, but also as chief architect and product manager. In addition, he has built up five startups in fields ranging from logistics to Internet-based services.

His research interests include requirements engineering, technology and product management, process assessment and improvement, quality assurance, and practical innovation. He is a member of IEEE and ACM.




Kai Petersen is a professor of software engineering at Blekinge Institute of Technology and works as a senior researcher at Ericsson AB. His research interests are in agile and lean software development, software testing, software process improvement, software security, and software measurement.



Michael Mattsson is a professor of software engineering at Blekinge Institute of Technology and Project Manager for the ORION project. His research interests include software architecture and component frameworks as well as decision-making and product line engineering.

Affiliations

Emil Alégroth¹  · **Tony Gorschek¹** · **Kai Petersen¹** · **Michael Mattsson¹**

Tony Gorschek
Tony.Gorschek@Bth.se

Kai Petersen
Kai.Petersen@Bth.se

Michael Mattsson
Michael.Mattsson@Bth.se

¹ Blekinge Institute of Technology, SE-371 79 Karlskrona, Sweden