Hird, Abigail and Mendibil, Kepa and Duffy, Alexander and Whitfield, Robert Ian (2015) New product development resource forecasting. R&D Management. ISSN 0033-6807 , http://dx.doi.org/10.1111/radm.12140

New Product Development Resource Forecasting

Abstract

Forecasting resource requirements for New Product Development (NPD) projects is essential for both strategic and tactical planning (Anderson Jr and Joglekar, 2005). Despite resource being the essential core of a business, priority is usually given to generating product and market intelligence whilst resource information is side-lined (Wernerfelt, 1984, 1995). This paper demonstrates that very little has changed nearly 30 years on from Wernerfelt's original work. Sophisticated, elegant planning tools to present data and inform decision making do exist (Kavadias and Loch, 2004; Kerzner, 2006). However, in NPD such tools run on unreliable, estimation-based resource information derived through undefined processes (Hird, 2013).

This paper establishes that existing methods do not provide transparent, consistent, timely or accurate resource planning information, highlighting the need for a new approach to resource forecasting, specifically in the field of NPD. The gap between the practical issues and available methods highlights the possibility of developing a novel Design of Experiments approach to create resource forecasting models.

Keywords

Resource forecasting, NPD, predictive modelling, estimation

1.0 Introduction

The process of planning the development of new products is fraught with uncertainty and complexity (Chalupnik et al., 2008; De Weck and Eckert, 2007; Joglekar et al., 2007; Sicotte and Bourgault, 2008). Novelty and innovation require exploration of the unknown, making some degree of uncertainty inherent and essential. The outcome of each design activity is unknown a priori. Consequently, the necessary proceeding activities cannot be known with certainty up-front (Haffey, 2007).

The New Product Development environment is one in which forecasting resource demand is particularly challenging (Anderson Jr and Joglekar, 2005; Loch and Terwiesch, 2007; Loch and Terwiesch, 1998). In most environments the goal of planning is to reduce uncertainty about events and their outcomes. Inhibiting uncertainty in NPD narrows the potential for innovation, defeating the objective of developing something new. However, not everything is uncertain and assumptions can be applied to the main types of activities that will be required and their likely outcomes (Kerzner, 2006). The problem of prediction is a complex one: multiple activities with multiple potential outcomes dictate proceeding activities; and, multiple factors can impact the likelihood of each outcome. Irrespective of the sophisticated planning tools that the resource data is packaged in, using an estimation-based approach to generate resource forecasts results in a number of issues (Hird 2013):

1. *Consistency.* Forecasts are formed through fluctuating perceptions, experience and judgements unique to each individual estimator (Yassine *et al.*, 2003). Apart from clouding the confidence that can be placed in predictions, this approach allows room for personal agendas and biases to

exist despite the potential for incompatibility with the best interests of the organisation (Ford and Sterman, 2003a, b).

2. *Transparency.* With estimations or unstructured analogies, the factors influencing the estimates are often tacit rather than explicit making it difficult for businesses to learn more about the considerations and processes that produce quality estimations. Personal agendas and biases can be perceived to exist within the estimation process which can create frustrations and lead to a whole new set of productivity issues.
3. *Timeliness.* Collecting, checking and updating estimations can be a lengthy, resource intensive process. By the time managers have been surveyed (across portfolio and functions) and the results compiled, estimates are often no longer valid as the scope of projects or portfolio of current projects may have changed. Even using a bottom-up approach, forecasts stored in plans must go through a cycle of checks and may need to be refreshed.
4. *Accuracy.* Unless accurate post-event data is recorded, the accuracy of forecasts will certainly remain unknown. Even with post-event recording (i.e. timesheets) evaluation of accuracy is not always clear-cut. For confidence to be placed in resource information and the decisions it informs, the accuracy must be known in the first instance. Lack of consistency, transparency and timeliness impact confidence in information accuracy.

Anderson and Joglekar (2005) present a hierarchical planning framework illustrating the cardinal role of resource information in NPD planning (See Figure 1). With such a pivotal role to play, are estimation-based forecasts the only suitable means of prediction in NPD? Our initial research question is framed: What techniques could be used for resource information generation in NPD?

[Please insert Figure 1 here]

Anderson and Joglekar (2005) do not describe where the resource information comes from other than from "raw operational data". However, what exactly "raw operational data" is and how it is derived is not clear. Assuming that "raw operational data" is resource information recorded post event, it is worth noting that this does not readily exist in NPD in practice. Gathering such data can be controversial: using timesheets is thought to inhibit productivity (Pfeffer and DeVoe, 2009; Webb, 1992); metrics developed through timesheets could conflict with forecast resource requirements exposing the lack of accuracy in original estimates; and, the measurements themselves could be unreliable (Pawar and Driva, 1999). The collective effort required to properly implement and manage the adoption of a robust resource information data-collection system inhibits widespread use of this approach in NPD planning processes.

The aim of this paper is to present the issues associated with the development of resource information in NPD and to explore the feasibility of alternative approaches. In order to achieve this a systematic literature review has been carried out. A large percentage of results from the initial review related to resource forecasting in the field of Software Development (SD). This prompted a second review focusing on resource forecasting methods specific to the field of SD. The methodology for both reviews is described (Section 2.0) followed by results from the NPD specific review (Section 3.0) and the results of the second SD specific review (Section 4.0). In Section 5.0, the feasibility of adapting an SD approach for NPD resource forecasting is discussed. The paper concludes by summarising the suitability of existing methods (Section 6.0). The gap between the practical requirements and the capabilities of existing methods leads to the proposal of a viable solution based upon a novel adaptation of Design of Experiments methodology described in Section 7.0. The need to verify the novel approach through case-study research and other avenues for future work are also discussed in Section 7.0.

2. 0 Literature review methodology

Both literature reviews follow the methodology for systematic reviews described by Tranfield et al., (2003). In both instances, two databases were used: ABI INFORM and Science Direct with a view to providing both business-focused and technically oriented results.

2.1 NPD focused literature review methodology

The objective of this first review was to qualify the attention generation of resource information receives in literature and to gain insight into the process of generating resource information in NPD. The abstracts of peer reviewed papers were searched using the following terms: "resource estimation OR "resource planning" OR "resource prediction" OR "resource forecasting" AND "NPD" OR "Product Development". Only nine papers were returned from ABI INFORM and 24 papers from Science Direct.

Results were reviewed by a single researcher with a view to establishing whether they addressed the issue of forming resource forecasts or otherwise. In instances where the paper did not directly address forming resource forecasts, the paper was categorised according to its main focus e.g. SD efforts or business growth. Given that the number of papers was low and that the question and answers sought were unambiguous it was decided that one reviewer was sufficient and a panel of reviewers was unnecessary.

A significant percentage of results from this first review referred to resource forecasting in the domain of Software Development (SD) and consequently, prompted a second literature search discussed below.

2.2 Software development literature review methodology

The objective of the second review is to identify methods used to predict resource demand in the field of Software Development. From the initial review it was clear that predictive modelling is commonly used in this domain and that “cost” and “effort” are used synonymously with “resource”. Initial search terms chosen were: “Predictive model” AND (“Cost” OR “Effort” OR “Resource”). To focus the search on most current methods, results were restricted to peer reviewed journal articles published between Jan 2010 and Feb 2013. The assumption was made that historically significant predictive modelling methods would be mentioned in any results and would thereby be included in our study. ABI/INFORM returned 521 articles and Science Direct: 7,800. Results featured a wide range of topics from hospital policy to predictors of homelessness As this work is only interested in research relating to methods used to predict effort or resource or cost search terms were narrowed by the following categories pre-defined by ABI/INFORM: Planning; Product and Process Development; Research and Development; HR Planning; Software and Systems; Management Science and Operations Research and the following topic for Science Direct: Predictive Modelling; Software. 152 articles were returned from ABI/INFORM and 72 from Science Direct. The 222 articles were manually sorted to establish relevance. Of the 222 articles, five were found to relate specifically to resource forecasting in the field of Software Development. Figure 2 illustrates the search process.

[Please insert Figure 2 here]

Each of the relevant papers was coded in Nvivo in order to identify the variety of methods used to forecast resource requirements as well as their strengths and weaknesses.

3.0 NPD focused literature review results

Table 1 describes the 31 results in terms of the subject area covered and the contribution of the paper to this study. The majority of the papers relate to Enterprise Resource Planning (ERP) systems and ERP implementation. Of the 31 results, three refer to quantifying resource demand in NPD and one relates to quantifying resource demand in SD. Of the three papers that relate to NPD, one merely mentions that understanding resource availability is a critical success factor and the other two focus on managing overall capacity from a portfolio planning viewpoint: no mention is given to quantifying resource demand.

[Please insert table 1 here]

In addition to highlighting the shortage of research on the subject, the review brings several other points to our attention.

- The emphasis upon project related data as opposed to resource related data remains. For example one paper discusses planning mechanisms (Tripathy and Biswal, 2007) whilst another focuses on product related factors impacting project success (Khurana and Rosenthal, 1998). The papers that do refer to resource do not mention developing resource information, rather there is a focus on dividing the given resource pool across different types of product development projects (Cooper, 1987; Yu et al., 2010).

- Research in the field tends to focus upon organising and presenting the information without details of how the resource information is derived (Tripathy and Biswal, 2007; Beaujon et al., 2001; Chao and Kavadias, 2008; Cooper et al., 1999; Kavadias and Loch, 2004; Loch and

Kavadias, 2002). Consistently, the focus is on the mechanism for portfolio management rather than the quality of the information feeding the mechanism.

- When resource information is considered, there is a tendency to focus on very general heuristics for high level capacity considerations rather than quantified resource (Cooper, 1987; Yu et al., 2010).

- The domain of SD has adopted a more scientific, evidence-based approach to generating resource forecasts (Subramanyam et al., 2012).

The low number of search results is surprising given the critical nature of resource prediction to NPD planning discussed by Anderson and Joglekar (2005). No obvious solution to the NPD resource forecasting issue exists. An intuitive, estimation-based approach remains the only solution to this complex issue. The only indication of an approach that offers an alternative is specific to the field of SD (Subramanyam et al., 2012). This paper contains reference to extensive work carried out in the SD specific domain and resulted in further research questions: which methods are used to forecast resource demand in SD? Can such methods be useful in alleviating the issues in NPD resource planning?

4.0 SD focused literature review results

The first objective of the review is to establish the range and nature of methods used in SD. Once this has been established, consideration is given to the possibility of applying software methods to NPD. In order to assess the suitability of each method the key difficulties are used as evaluation criteria. These are: the accuracy of the method, the consistency of the method, the transparency of the method and the time required to generate and verify estimations (Hird, 2013).

4.1 SD resource forecasting methods

Each of the five papers from the literature search features a variety of different forecasting methods and are illustrated in Table 2. Although each paper contains aspects of comparison between methods, it is difficult to establish conclusively which methods are superior. Apart from different contexts requiring different approaches, attempts at meta-analysis are hindered by the use of disparate data sets, analysis methods and evaluation criteria (See Dejaeger et al., 2012; Kitchenham et al., 2007; Wen et al., 2012).

SD resource forecasting methods can be divided into three groups: estimation based, theoretical, and historical data based. The methods described in each paper are presented in Table 2 below before they are discussed in more detail and related to the criteria presented by Hird (2013).

[Please insert Table 2 here]

4.2 SD forecasting method suitability

In this section, each of the forecasting methods identified in Table 2 is discussed. Discussion has been grouped according to the three types of method: estimation based, theoretical, and historical data based methods.

4.2.1 Estimation-based methods

The estimation-based or human-centric approach to generating resource demand predictions is still widely used in software development (Dejaeger et al., 2012; Lederer and Prasad, 1993; Wen et al., 2012). With no suitable alternative, it is certainly the most widely used method in companies that develop non-software products. The fundamental issue with estimation-based methods is the lack of objectivity coupled with reliance upon domain experts (Dejaeger et al., 2012). Biases and personal agendas are accommodated within estimation-based planning, placing power in the hands of resource owners (who may have a narrow, localised view) as opposed to a perspective and agenda aligned with wider business goals.

Although there are specific instances where an expert is likely to be more accurate (Dejaegar et al, 2012) there are situations where models can allow reduction in situational and human biases. Lederer and Prasad (1993) found unaided-estimates to be inaccurate, demonstrating a positive correlation between the use of intuition and the percentage of projects overrunning their estimates.

12 best practice guidelines for expert estimation are evaluated and endorsed by Jørgensen (2004). Although the guidelines are excellent in theory, the "how" for implementing the guidelines in practice remains elusive. For example, how can "conflicting estimation goals" be avoided and how can we determine which information is "irrelevant and unreliable"?

From the literature search conducted, DELPHI was the only expert estimation technique specifically mentioned (Dejaeger et al., 2012). DELPHI involves several domain experts formulating independent estimates which are subsequently collated. Either the median is used as the final effort estimation or the process can iterate: independent estimates may be anonymised and re-distributed to the panel of experts for further consideration, adjusted (if required), re-collated, the median re-calculated and the refined estimation used.

Employing the 12 best practice guidelines or methods such as DELPHI may well result in less ambiguous estimates but coordinating such activities requires considerable time and effort especially when estimations are required across a large number of projects and/or functional groups. By employing this method, the reliance upon domain experts remains and there is no flexibility for a portfolio manager to generate predictions quickly for the purposes of "what-if" scenario analysis.

4.2.2 Theory-based methods

Models bring consistency and the ability to correctly assess the impact of different inputs in a fashion that the limits of human experience and cognitive capabilities prohibit (Ayres, 2007). With the theory-based modelling method an expert proposes a general model then domain data are used to model specific projects. For example, the COnstructive COst MOdel (COCOMO) describes both a predictive algorithm and tuning procedure (Boehm, 1981). This tuning procedure requires significant effort from the business (Dejaeger et al., 2012) yet the model must be tuned if it is to be used effectively (Kitchenham et al., 2007). COCOMO I is generally regarded as out-dated and has been updated (Boehm, 2000). However, the data supporting the updated model is not yet publically available.

In the most recent literature, formal theory-based models are given diminishing attention as they are superseded by data-based models enabled by machine learning techniques. The main limitation of theory-based models is the association with a fixed, specific set of attributes. If projects cannot be described in line with the attributes then the model is rendered useless in that context (Dejaeger et al., 2012). Additionally, the information used to generate the prediction is often itself estimation-based especially early on in a project (El-Sebakhy et al., 2012).

4.2.3 Data-based methods

Data-based models are based upon the identification of trends in sets of historical project data. Such trends describe the relationship between numerous independent variables (project characteristics) and a dependant variable (resource, cost or effort requirements).

The main benefit of data-based models is the resultant objective, analytical process insights and predictions. Such an approach is recommended by ISO 9000 which states that "*effective decisions are based upon the analysis of data and information*". With data-based models there is flexibility to specify the most appropriate independent variables which makes them an attractive option (Dejaeger et al., 2012).

Broadly speaking, sourcing data to create predictive models of this type is the key limitation. Vast quantities of data are required to establish reliable data-based models (Jørgensen, 2004). Data requirements form a barrier making data-based models for development companies an unachievable option: the time required to accumulate enough data on past projects from a single company may be prohibitive; by the time the data set is large enough to be of use technologies used by the company may have changed and older projects may no longer be representative of current practices; and care is necessary as data needs to be collected in a consistent manner Kitchenham et al., (2007).

Although data based methods provide opportunity to develop accurate practical models, having reliable, accurate data is a prerequisite (Maxwell et al., 1999). Software companies are able to pool data as the attributes of software are consistent (lines of code, function points). Non-software development companies on the other hand would find less use in pooling data as each project is idiosyncratic.

Our literature search revealed several data-based predictive modelling methods. Some of the methods listed in Table 2 are extensions of other methods whilst some (for example MARS) have not actually been applied in practice and remain purely theoretical. Rather than delving into the technical details, in the following sections we review the key types of data-based modelling methods and the overriding strengths and limitations for each. The following types of method are reviewed: linear modelling methods, non-linear modelling methods and tree based methods.

Linear modelling methods

Regression modelling is one of the most widely applied techniques for software effort estimation. This well-documented technique fits a linear regression function to a data set containing a dependent variable (effort, cost, resource) and multiple independent variables (Dejaeger et al., 2012).

Not only is a large volume of historical data required to build a regression model, the data used ought to meet several assumptions (normality, linearity of relationships, independence of errors and homoscedasticity of the errors versus time and predictions). In practice, these assumptions are rarely adhered to strictly (Kitchenham, 1992); linear models are often compromised in their predictive abilities. Various statistical treatments can be applied to help bring the data set closer to the specified assumptions (Kocaguneli et al., 2012; Kitchenham and Mendes, 2009).

Two further limitations of linear modelling exist: the number of factors (project characteristics) that can be included in the model is limited by the size of the data set; and, the inclusion and identification of confounding variables. It is possible, if not probable that some of the factors will not impact resource demand independently. To establish whether or not this is the case, each possible interaction has to be tested. Confidently establishing which factors and interactions to include in the model is one limitation with this approach.

In practice, the vast quantities of data required; the constraints of the assumptions; and restrictions with regards to the modelling of interactions between variables render linear regression less than ideal for modelling the complex factors driving resource demand.

Non-linear models

Where linear models fall down with their ridged assumptions, non-linear models are capable of modelling or at least approximating any form of arbitrary distribution, normal or otherwise. Several types of non-linear methods exist and are referred to as Machine Learning (ML) models. Non-linear models fall into two main groups: Artificial Neural Networks (ANNs) and Tree and Rule based models. More obscure categories exist for example Genetic Programming (Burgess and Lefley, 2001), Fuzzy-logic models (Xu and Zhang, 2012) and Bayesian statistics (Chulani et al., 1999). The more obscure methods are usually initiated by or integrated within ANNs and can be described by the same benefits and limitations.

Artificial Neural Networks

Artificial neural networks are adept to modelling the nuances of complex systems. They are capable of approximately representing arbitrary distributions and non-parametric relationships which is useful as data is rarely distributed in linear functions (Witten et al., 2011; Yegnanarayana, 2004). This learning ability of ANNs is particularly significant if the factors affecting resource demand are likely to vary over time (Maxwell et al., 1999).

One of the major drawbacks of ANNs in this context is the lack of transparency within the underlying model; the nature of the relationships between the input and output cannot be inferred. Additionally, relatively large volumes of data are required to train and validate the network.

Tree and Rule based models

Tree and Rule based models are a ML technique with very different properties to ANN. Rather than learning or training a model, the actual, historical project data is organised and represented in a tree like structure. Predictions could be made for any future project with the same characteristics as an existing project by following the structure of the tree.

The comprehensibility of regression trees can be considered one of the main strengths of this technique along with the explicit ability to model interactions and arbitrary distributions. The limitations of this technique are the ability to extrapolate for cases that are not represented by historical projects and the categorical rather than continuous nature of the outputs (Witten et al., 2011). Case Based Reasoning or CBS is a less sophisticated version where analogous projects are sought without the informative structure (Mukhopadhyay et al., 1992).

5.0 Developing a predictive modelling approach for NPD

Regardless of whether it is applied in practice, the process of developing a model and considering the factors impacting resource demand should, in itself, be a worthwhile endeavour: useful business insights can be generated (Subramanyam et al., 2012).To establish the applicability of the SD methods to NPD, the fundamental differences are discussed. Owing to the unique nature of NPD and the limitations of the context, additional predictive modelling requirements are proposed before each of the SD methods is considered in terms of its potential to meet the requirements.

5.1 Differences between Software Development and NPD

The majority of Software Development resource estimation methods are based upon data and statistical analysis or machine learning. Even theory based methods are based upon trends in data or general heuristics derived from regression. When compared with NPD, Software Development effort estimation is more suited to a predictive modelling for three main reasons.

Firstly, software projects can be characterised in similar ways regardless of the field of application (for example, lines of code or function points) whereas the characteristics of projects which don't exclusively feature software are not so easy to group. From naval ships and aircraft to medical devices, consumer products or automobiles: each grouping has distinct characteristics. Theory-based models are based upon commonality and general applicability. A general model can be adapted, tuned and tailored to the needs of specific software development companies. The generic nature of such models and the changing face of software development have instigated the need for updates, increasing adaptions and more laborious tuning (Boehm, 2000; Dejaeger et al., 2012).

Secondly, over the past decades, pools of historical data have been collected either by organisations such as International Software Benchmarking Standards Group (ISBSG), between groups of companies or within specific companies. Simultaneously, data mining techniques have improved, allowing researchers to explore inductive, non-generic predictive models specific to data sets.

Thirdly, it is possible that the notion of data driven solutions to complex problems and abstract concepts such as ML sit more comfortably and are more readily accepted by the software community who routinely deal with the virtual and intangible compared to hardware product developers who are more familiar with the physical and visible.

In addition to the considerations set out in the introductory section (consistency, transparency, timeliness and accuracy), a further consideration is the capability of the predictive modelling methods to consider confounding variables. The ability to model confounding variables will be included as one of the requirements for an appropriate resource forecasting method. Each of the methods is considered against the requirements in Table 3 below.

[Please insert Table 3 here]

6.0 Discussion: Steps towards improving resource planning

None of the existing methods provide a solution suitable for application in an NPD environment which isn't SD. The main issue is the lack of historical project data. If suitable data was available upon which a model could be constructed and verified then it is possible that the lack of transparency of ANN or the inability of linear regression to model confounding variables could be accepted. It is clear why, with no other realistic option, estimates are used so prolifically despite their inadequacies.

From a logical perspective the regression tree holds the most potential: it can be accurate, consistent and is very transparent in nature. The only issue with such a method is the missing data. Estimations could be made for the "missing" scenarios but with so many missing scenarios, the obvious question is how would we know which scenarios to look at? It would take an unreasonable and perhaps impossible amount of time to conduct a survey to gather estimates for every possible project type.

Is there a systematic means of identifying which scenarios or branches of the regression tree could provide us with the most complete model in the most efficient manner possible?

Although not previously discussed within resource forecasting literature, a method from another field exists which could help identify a solution to the question posed above. Within the field of quality control there is a popular method that allows a series of experiments to be designed in the most efficient way possible. The Design of Experiments (DoE) method specifies an economic series of experiments that will return results which allow maximum insight into factors driving variation. DoE, when carried out in full, allows the creation of a predictive regression model which describes the experimental variables under study. Rather than conducting an exhaustive series of experiments by altering one variable at a time, DoE allows approximate relationships to be determined from a minimal number of experimental runs.

DoE has always been used when experiments can be carefully controlled *a priori* and results can be measured. Such a method could be used to identify the key project scenarios (from potentially hundreds of thousands) that would allow us to gain insight to the factors driving resource demand. Rather than gathering estimations for each and every possible scenario. It is possible that a method such a DoE could be applied to create a view of the relationships between factors impacting resource demand and resource demand through consideration of a few carefully designed projects. Through utilising estimations instead of experimental results and hypothetical project scenarios instead of designed experiments, it is possible that a practical model could be constructed. This model could potentially provide accuracy as estimates would be about hypothetical rather than actual scenarios removing biases based on “resource ownership” concerns. We envisage that such a model could be constructed in a “one-off” effort using carefully considered estimation perhaps in-line with the suggestions proposed by Jørgensen, (2004) or using the DELPHI technique.

An outline of the proposed novel process is included in Figure 3 below.

[Please insert figure 3]

It is possible that such a method could meet the criteria specified. Such a model could foreseeably be developed from the resulting regression equations to provide consistent and timely resource forecasts. If hypothetical project scenarios were used in conjunction with estimations, no historical data would be required and the factors impacting resource demand would be transparent and easily identified. As the technique is based upon estimations, the ability of the model to provide accurate forecasts is not certain and requires further work.

The problems faced with regards to resource planning may not be peculiar to NPD. Research should be conducted to explore other domains which could benefit from such an approach. Such domains are likely to also be characterised by high levels of uncertainty and complexity for example the allocation of humanitarian aid and disaster relief resources or cost modelling for products based on new technologies

Application of such a technique could also offer a means of assessing the accuracy and consistency of the estimations made by planners, either for the purposes of a performance measurement tool or in order to create a method of expert knowledge capture. Application of a

DoE approach could also provide insights into the estimation process and could allow existing heuristics to be tested or allow new evidence-based heuristics to be derived

7.0 Conclusions and future work

This paper demonstrates inconsistency between the importance of resource forecasting in NPD and the attention it is given by researchers. A key contribution from this paper is evidence suggesting that existing methods do not provide transparent, consistent, timely or accurate resource planning information. This work contributes to the field of NPD resource planning by demonstrating that currently, an estimation based approach is the only available option for forecasting resource and more research is required to develop forecasting techniques that are meet practical requirements.

As a result of the contribution made, we propose further investigation into a new combination approach based upon the novel application of DoE and a formalisation of the estimation process which has potential to result in a timely, consistent and transparent model which would be capable of making interactions between variables explicit and would not require historical data (other than for validation purposes). Such a technique, should it be proven viable could potentially be used in conjunction with other methods. It could be used to generate training data for ANN or could be used to provide upfront information about which interactions exist and which factors to include in regression models.

Providing the capability to forecast resource demand is a fundamental aspect of resource planning. As other aspects of resource planning rely upon this resource demand information, improvements in this area could open the doors to improving other aspects of resource planning for example: improved levelling of resource capacity and improved resource allocation in portfolio optimisation.

References

Anderson Jr, E. G., and N. R. Joglekar, 2005, A hierarchical product development planning framework: Production and Operations Management, v. 14, p. 344-361.

Ayres, I., 2007, Super crunchers: how anything can be predicted, John Murray.

Beaujon, G. J., S. P.Marin, and G. C. McDonald, 2001, Balancing and optimizing a portfolio of R&D projects: Naval Research Logistics (NRL), v. 48, p. 18-40.

Boehm, B. W., 1981, Software Engineering Economics, Pearson Education.

Boehm, B. W., 2000, Software cost estimation with Cocomo II, Prentice Hall.

Burgess, C. J., and M. Lefley, 2001, Can genetic programming improve software effort estimation? A comparative evaluation: Information and Software Technology, v. 43, p. 863-873.

Castejón-limas, M., J. Ordieres-meré, A. González-marcos, and V. González-castro, 2011, Effort estimates through project complexity: Annals of Operations Research, v. 186, p. 395-406.

Chalupnik, M. J., D. C. Wynn, C. M. Eckert, and P. J. Clarkson, 2008, Investigating design process performance under uncertainty.

Chao, R., and S. Kavadias, 2008, A theoretical framework for managing the NPD portfolio: when and how to use strategic buckets: Management Science, Vol. 54, No. 5, pp. 907-921, 2008.

Cheung, W. M., P. G. Maropoulos, and P. C. Matthews, 2010, Linking design and manufacturing domains via web-based and enterprise integration technologies: International Journal of Computer Applications in Technology, v. 37, p. 182.

Chien, S.-W., C. Hu,K. Reimers, and J.-S. Lin, 2007, The influence of centrifugal and centripetal forces on ERP project success in small and medium-sized enterprises in China and Taiwan: International Journal of Production Economics, v. 107, p. 380-396.

Chryssolouris, G., D. Mavrikios, N. Papakostas, D. Mourtzis, G. Michalos, and K. Georgoulias, 2009, Digital manufacturing: history, perspectives, and outlook: Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, v. 223, p. 451-462.

Chulani, S., B. Boehm, and B. Steece, 1999, Bayesian analysis of empirical software engineering cost models: Software Engineering, IEEE Transactions on, v. 25, p. 573-583.

Colvin, M., and C. T. Maravelias, 2009, Scheduling of testing tasks and resource planning in new product development using stochastic programming: Computers & Chemical Engineering, v. 33, p. 964-976.

Cooper, R. G., 1987, Defining the New Product Strategy: IEEE Transactions on Engineering Management, v. EM34, p. 184-184.

Cooper, R. G., S. J. Edgett, and E. J. Kleinschmidt, 1999, New product portfolio management: practices and performance: Journal of product innovation management, v. 16, p. 333-351.

Davila, A., and G. Foster, 2007, Management Control Systems in Early-Stage Startup Companies: The Accounting Review, v. 82, p. 907-937.

De Weck, O., and C. Eckert, 2007, A Classification of Uncertainty for Early Product and System Design.

Dejaeger, K., W. Verbeke, D. Martens, and B. Baesens, 2012, Data Mining Techniques for Software Effort Estimation: A Comparative Study: IEEE Transactions on Software Engineering, v. 38, p. 375-397.
Ding, Y., and B. Sheng, 2010, ERP and PDM integration technology to support collaborative product development: International Journal of Business Information Systems, v. 5, p. 440-459.
Duffy, V. G., and G. Salvendy, 2000, Concurrent engineering and virtual reality for human resource planning: Computers in Industry, v. 42, p. 109-125.
El-Sebakhy, E. A., O. Asparouhov, A.-A. Abdulraheem, A.-A. Al-Majed, D. Wu, K. Latinski, and I. Raharja, 2012, Functional networks as a new data mining predictive paradigm to predict permeability in a carbonate reservoir: Expert Systems with Applications, v. 39, p. 10359-10375.
Feng, S. C., 2000, Manufacturing planning and execution software interfaces: Journal of Manufacturing Systems, v. 19, p. 1-17.
Finnie, G., G. Wittig, and J. Desharnais, 1997, A comparison of software effort estimation techniques: using function points with neural networks, case-based reasoning and regression models: Journal of Systems and Software, v. 39, p. 281-289.
Ford, D. N., and J. D. Sterman, 2003a, Overcoming the 90% syndrome: Iteration management in concurrent development projects: Concurrent Engineering, v. 11, p. 177.
Ford, D. N., and J. D. Sterman, 2003b, The Liar's Club: concealing rework in concurrent development: Concurrent Engineering, v. 11, p. 211.
Fortin, C., and G. Huet, 2007, Manufacturing Process Management: iterative synchronisation of engineering data with manufacturing realities: International Journal of Product Development, v. 4, p. 280-280.
France, K., 2002, The future of information technology in collaborative construction: Civil Engineering, v. 150, p. 4-6.
Gao, J. X., H. Aziz, P. G. Maropoulos, and W. M. Cheung, 2003, Application of product data management technologies for enterprise integration: International Journal of Computer Integrated Manufacturing, v. 16, p. 491-500.
Goossenaerts, J. B. M., A. T. M. Zegers, and J. M. Smits, 2009, A multi-level model-driven regime for value-added tax compliance in ERP systems: Computers in Industry, v. 60, p. 709-727.
Guo, X.-J., T.-Y. Wang, X.-D. Qin, C.-E. Wang, and H.-Y. Li, 2005, Research on J2EE-based enterprise integrated platform & application system for electric equipment enterprise: Jisuanji Jicheng Zhizao Xitong/Computer Integrated Manufacturing System (China), v. 11, p. 927-931,995.
Haffey, M., 2007, An approach, insights and Methodology for perfromance improvement through process activity management, Univerity of Strathclyde, Glasgow, 597 p.
Hird, A., 2013, A Systems Approach to Resource Planning in New Product Development, University of Strathclyde.
Hou, L., D. Han, and Z. Lin, 2006, Sourcing configuration for product mass customization: Jixie Gongcheng Xuebao (Chinese Journal of Mechanical Engineering), v. 42, p. 131-136.
Hurtarte, J. S., E. A. Wolsheimer, and L. M. Tafoya, 2007, Chapter 11 - Best Practices for Fabless Companies, Understanding Fabless IC Technology: Burlington, Newnes, p. 163-184.

Joglekar, N. R., N. Kulatilaka, and E. G. Anderson Jr, 2007, 11 Hierarchical planning under uncertainty: Real options and heuristics: Handbook of new product development management, p. 291.
Jørgensen, M., 2004, A review of studies on expert estimation of software development effort: Journal of Systems and Software, v. 70, p. 37-60.
Kavadias, S., and C. H. Loch, 2004, Project selection under uncertainty: dynamically allocating resources to maximize value, Springer Netherlands.
Kerzner, H., 2006, Project management: a systems approach to planning, scheduling, and controlling, Wiley.
Khurana, A., and S. R. Rosenthal, 1998, Towards holistic "front ends" in new product development: Journal of Product Innovation Management, v. 15, p. 57-74.
Kitchenham, B., and E. Mendes, 2009, Why comparative effort prediction studies may be invalid: Proceedings of the 5th International Conference on Predictor Models in Software Engineering, p. 4.
Kitchenham, B. A., 1992, Empirical studies of assumptions that underlie software cost-estimation models: Information and Software Technology, v. 34, p. 211-218.
Kitchenham, B. A., E. Mendes, and G. H. Travassos, 2007, Cross versus Within-Company Cost Estimation Studies: A Systematic Review: IEEE Transactions on Software Engineering, v. 33, p. 316-316.
Kocaguneli, E., T. Menzies, A. Bener, and J. W. Keung, 2012, Exploiting the essential assumptions of analogy-based effort estimation: Software Engineering, IEEE Transactions on, v. 38, p. 425-438.
Lederer, A. L., and J. Prasad, 1993, Information systems software cost estimating: a current assessment: Journal of information technology, v. 8, p. 22-33.
Lehrer, M., and M. Behnam, 2009, Modularity vs programmability in design of international products: Beyond the standardization-adaptation tradeoff?: European Management Journal, v. 27, p. 281.
Liew, C.-B. A., 2008, Strategic integration of knowledge management and customer relationship management: Journal of Knowledge Management, v. 12, p. 131-146.
Loch, C., and C. Terwiesch, 2007, Coordination and information exchange. C. Loch, S. Kavadias, eds., Handbook of New Product Development Management: Butterworth Heinemann, Amsterdam, v. 315, p. 343.
Loch, C. H., and S. Kavadias, 2002, Dynamic portfolio selection of NPD programs using marginal returns: Management Science, p. 1227-1241.
Loch, C. H., and C. Terwiesch, 1998, Communication and uncertainty in concurrent engineering: Management Science, p. 1032-1048.
Maxwell, K., W. Luk Van, and S. Dutta, 1999, Performance evaluation of general and company specific models in software development effort estimation: Management Science, v. 45, p. 787-803.
Melnyk, S. A., and R. F. Gonzalez, 1985, MRP II: The Early Returns Are In: Production and Inventory Management, v. 26, p. 124-124.
Mukhopadhyay, T., S. S. Vicinanza, and M. J. Prietula, 1992, Examining the feasibility of a case-based reasoning model for software effort estimation: MIS quarterly, p. 155-171.
Paviot, T., V. Cheutet, and S. Lamouri, 2011, A PLCS framework for PDM/ERP interoperability: International Journal of Product Lifecycle Management, v. 5, p. 295-313.

Pawar, K. S., and H. Driva, 1999, Performance measurement for product design and development in a manufacturing environment: International journal of production economics, v. 60, p. 61-68.

Pendharkar, P. C., 2010, Probabilistic estimation of software size and effort: Expert Systems with Applications, v. 37, p. 4435-4440.

Pfeffer, J., and S. E. DeVoe, 2009, Economic evaluation: The effect of money and economics on attitudes about volunteering: Journal of Economic Psychology, v. 30, p. 500-508.

Riel, A., S. Tichkiewitch, and R. Messnarz, 2010, Qualification and certification for the competitive edge in Integrated Design: CIRP Journal of Manufacturing Science and Technology, v. 2, p. 279-289.

Sicotte, H., and M. Bourgault, 2008, Dimensions of uncertainty and their moderating effect on new product development project performance: R&D Management, v. 38, p. 468-479.

Strehle, F., B. R. Katzy, and T. Davila, 2010, Learning capabilities and the growth of technology-based new ventures: International Journal of Technology Management, v. 52, p. 26-45.

Subramanyam, R., N. Ramasubbu, and M. S. Krishnan, 2012, In Search of Efficient Flexibility: Effects of Software Component Granularity on Development Effort, Defects, and Customization Effort: Information Systems Research, v. 23, p. 787-803,845-847.

Tranfield, D., D. Denyer, and P. Smart, 2003, Towards a methodology for developing evidence-informed management knowledge by means of systematic review: British Journal of Management, v. 14, p. 207-222.

Tripathy, B. B., and M. P. Biswal, 2007, A zero-one goal programming approach for project selection: Journal of Information & Optimization Sciences, v. 28, p. 619-626.

Vilpola, I., 2008, Development and Evaluation of a Customer-Centred ERP Implementation Methods. JITTA : Journal of Information Technology Theory and Application, v. 9, p. 1-21.

Webb, J., 1992, Themismanagement of innovation: Sociology, v. 26, p. 471.

Wen, J., S. Li, Z. Lin, Y. Hu, and C. Huang, 2012, Systematic literature review of machine learning based software development effort estimation models: Information and Software Technology, v. 54, p. 41-59.

Wernerfelt, B., 1984, A resource-based view of the rm: Strategic Management Journal, v. 5, p. 171–180.

Wernerfelt, B., 1995, The resource based view of the firm: Ten years after: Strategic Management Journal, v. 16, p. 171-174.

Witten, I. H., E. Frank, G. Holmes, and M. A. Hall, 2011, Data Mining: Practical Machine Learning Tools and Techniques, Elsevier Science.

Xu, J., and Z. Zhang, 2012, A fuzzy random resource-constrained scheduling model with multiple projects and its application to a working procedure in a large-scale water conservancy and hydropower construction project: Journal of Scheduling, v. 15, p. 253-272.

Yassine, A., N. Joglekar, D. Braha, S. Eppinger, and D. Whitney, 2003, Information hiding in product development: The design churn effect: Research in Engineering Design, v. 14, p. 145-161.

Yegnanarayana, B., 2004, Artificial neural networks, PHI Learning Pvt. Ltd.

Yu, A. S. O., P. S. Figueiredo, and P. T. d. S. Nascimento, 2010, Development Resource Planning:Complexity of Product Development and the Capacity to Launch New Products: The Journal of Product Innovation Management, v. 27, p. 253.

Zhao, S., and M. Yin, 2007, Research on an Information Integration Framework on the Large Complex Product Development PMIS with ERP: Materials Science Forum, v. 532-533, p. 917-920.

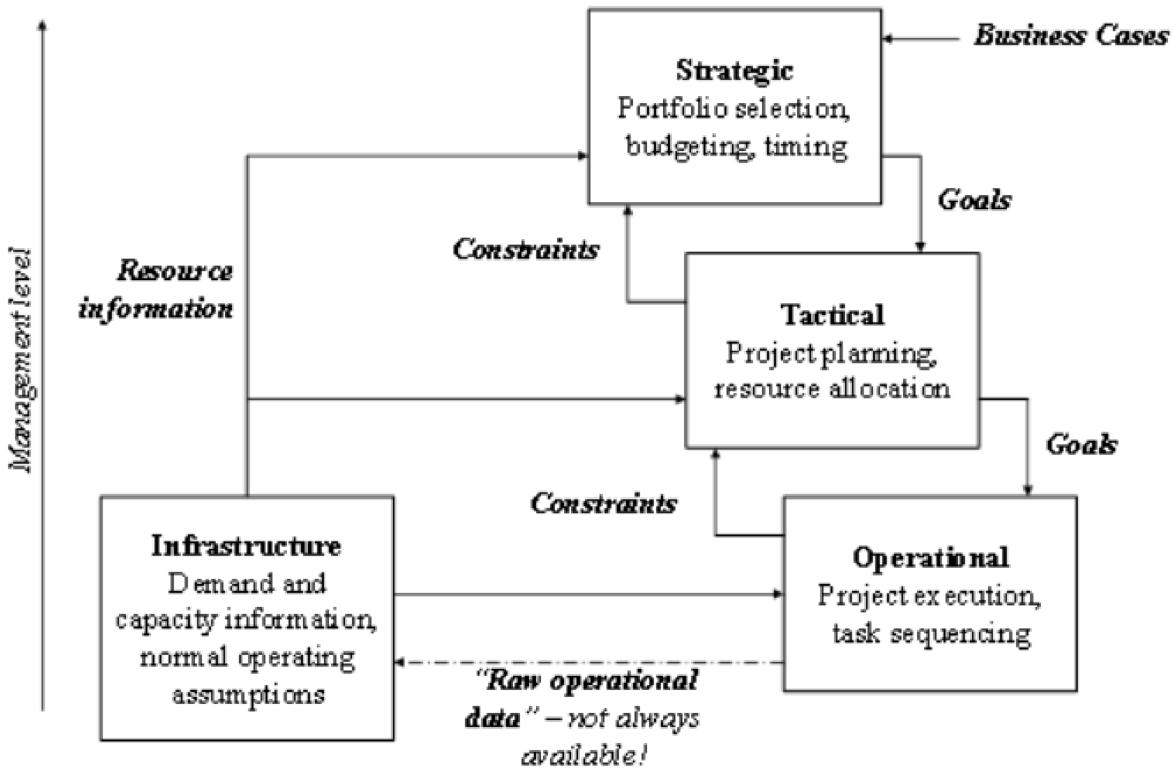

Figure 1 - Adaptation of Anderson and Joglekars (2005) Hierarchical planning framework

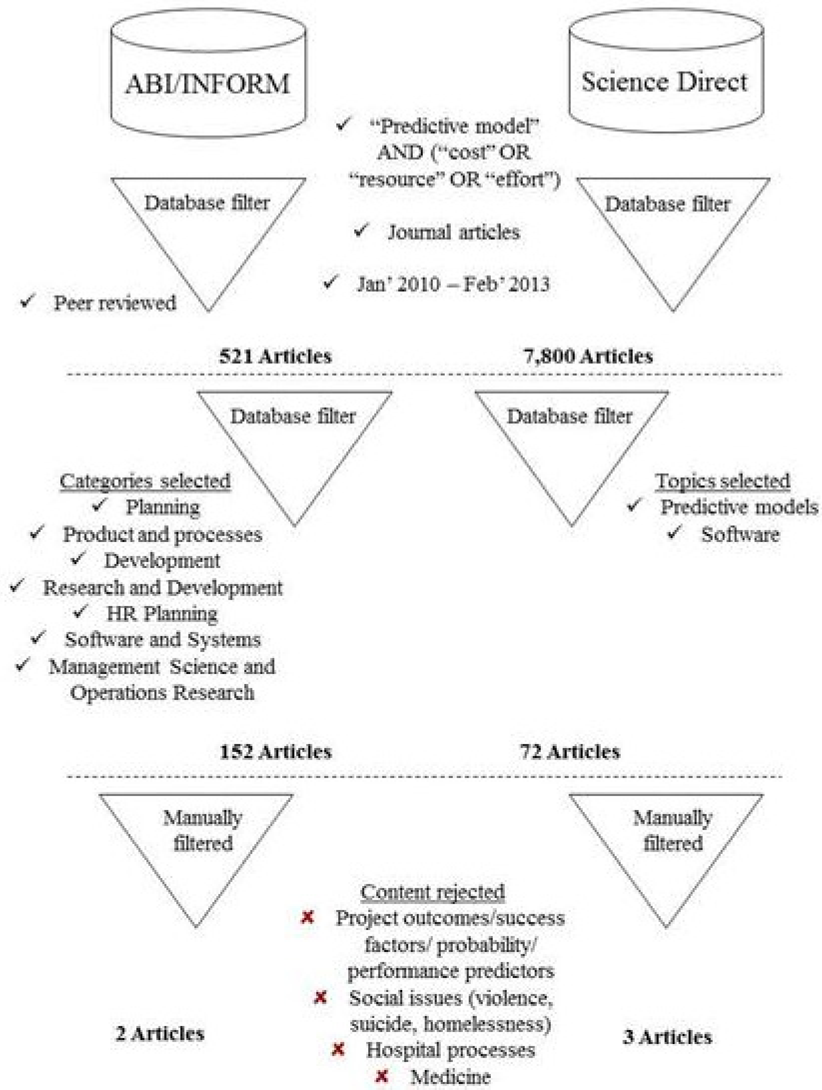

Figure 2 – SD specific review methodology

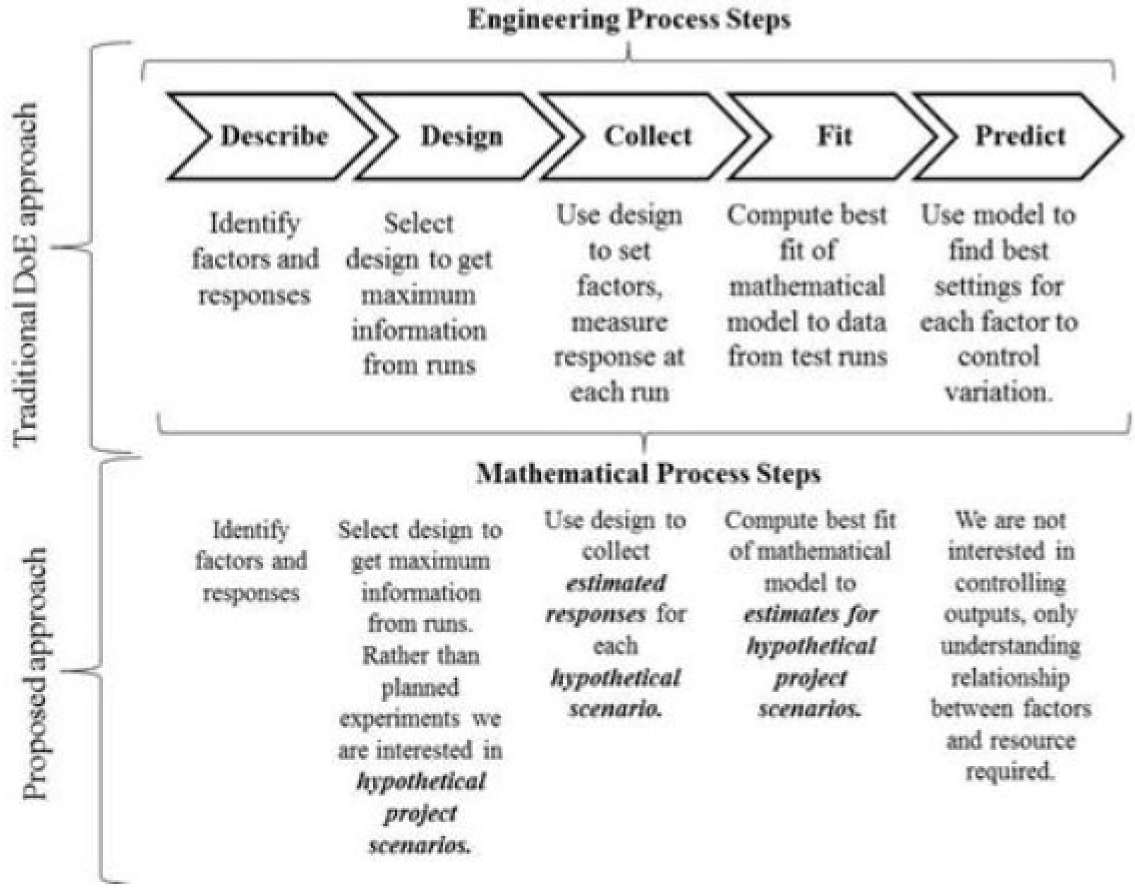

Figure 3 – Traditional Design of Experiments approach and novel proposed approach for forecasting resource requirements.

Table 1 – Literature search results

Subject area	References	Discussion of resource forecast generation
Front end processes. Factors impacting success.	Khurana and Rosenthal, 1998	Balance of risks and resource availability is cited as a risk factor although no detail is provided.
Resource planning in NPD	Cooper, 1987; Yu et al., 2010	Focus on development capacity rather than resource demand.
Software development effort	Subramanyam et al., 2012	Regression analysis used to examine links between the characteristics of the product and development effort / efficiency.

Scheduling using stochastic programming	Colvin and Maravelias, 2009	Tools exist to manage and present and manipulate data to aid decision making although the fundamental information driving decision making may be flawed.
Program selection, decision making process	Tripathy and Biswal, 2007	
Scheduling using stochastic programming	Colvin and Maravelias, 2009	None
Product standardisation /complexity /customisation.	Hou et al., 2006; Lehrer and Behnam, 2009	
Certification and qualifications	Riel et al., 2010	
Virtual reality communication tools	Duffy and Salvendy, 2000	
Business growth and start-ups	Davila and Foster, 2007; Strehle et al., 2010	
Product and manufacturing data management tools	Feng, 2000; France, 2002; Gao et al., 2003; Melnyk and Gonzalez, 1985	
ERP success, implementation, systems etc.	Cheung et al., 2010; Chien et al., 2007; Chryssolouris et al., 2009; Ding and Sheng, 2010; Feng, 2000; Fortin and Huet, 2007; Goossenaerts et al., 2009; Guo et al., 2005; Hurtarte et al., 2007; Lehrer and Behnam, 2009; Liew, 2008; Paviot et al., 2011; Vilpola, 2008; Zhao and Yin, 2007	

Table 2: Software project effort prediction methods

Approach		Method type, acronyms.	Data mining techniques for Software Effort Estimation: A Comparative Study (Dejaeger et al., 2012)	Functional Networks as a novel data mining paradigm in forecasting software development efforts (El-Sebakhy et al., 2012)	Probabilistic estimation of Software Size and Effort (Pendharkar, 2010))	A Systematic review of machine learning based software development effort estimation models (Wen et al., 2012)	Effort Estimates through project complexity (Castejón-limas et al., 2011)
Data-based method; (induced prediction system)		Linear modelling	Multiple regression	Regression	Regression		Linear regression
	Machine learning methods, non-parametric techniques	Non-linear modelling methods	ANN MARS Least-squares vector	ANN Neuro-fuzzy logic inference systems Bayesian statistics	Bayesian networks Genetic algorithms Genetic programming Rule Induction *CHAID* *PNN*	ANN Bayesian Networks Support Vector regression Genetic algorithms Genetic programming Association rules Fuzzy logic	ANN
		Tree or rule based methods; Classification techniques	CART	CBR (Analogy based estimation)	CART CBR	CBR Decision Trees	
Theory based method		Formal models; model based		COCOMO	COCOMO SLIM	COCOMO SLIM	
				Function Point Analysis		*Function Point Analysis*	
Estimation based method			DELPHI			Expert judgement	

Table 3 – Methods versus requirements for use in NPD context

Method	Transparency	Consistency	Timeliness	Confounding variables detected or accounted for?	Data requirements
Regression	Satisfactory	Stable.	Good.	Not detected but can be accounted for.	Large amounts of past project data required.
Artificial Neural Networks	Poor	Excellent: Semi-stable. Learns as parameters change.	Good.	Yes, both	Large amounts of past project data required.
CART	Good. The tree structure provides an explicit, traceable, visible and instinctive breakdown.	Stable: the same algorithm is used every time. Or semi-stable: new data can be added as it become available.	This method is timely providing the project characteristics are familiar, otherwise no result will be returned.	Yes, both	Large amounts of past project data required for comprehensive modelling. Although small data sets can also be modelled.
Formal models	Good	Stable	Timely once established. Reasonable knowledge of project scope required.	No, neither	Some data required to tune model. Project must have clear scope before this method can be applied.
Estimation with Delphi	Poor - Good.	Poor	Poor	Yes	Good, no historical data required.
Unstructured estimation	Poor	Poor	Poor	Yes	Good, no historical data required.