



# The combined impact of stochastic and correlated activity durations and design uncertainty on project plans

Michal Kaut<sup>a, \*\*</sup>, Hajnalka Vaagen<sup>b, \*</sup>, Stein W. Wallace<sup>c</sup>

<sup>a</sup> SINTEF Industry, Trondheim, Norway

<sup>b</sup> Norwegian University of Science and Technology (NTNU), Ålesund, Norway

<sup>c</sup> NHH Norwegian School of Economics, Bergen, Norway

## ARTICLE INFO

### Keywords:

Project planning  
Design uncertainty  
Correlations  
Flexibility  
Learning behaviour

## ABSTRACT

Most model based studies on project uncertainty investigate a single source of uncertainty, with a dominant focus on stochastic activity durations. However, another major uncertainty facing engineering projects is that of changes in design throughout the project delivery. This may come from uncertainty in the market, technology, or regulations, leading to changes in design and implementation paths, with alterations in the project network itself. This comes on top of stochastic and correlated activity durations for a given design. In this paper we develop a stochastic program to investigate how uncertainty in design and activity durations, together, affect planning, and their relationships. The findings suggest that when design uncertainty is modelled by multiple alternatives and delayed decisions on the final alternative, stochastic and correlated activity durations have limited impact. In situations with alternative and substitutable solutions available for a given design, correlations drive a certain learning behaviour.

## 1. Introduction

This paper treats project uncertainty and planning decision making in construction and engineering projects, where frequent changes in the scope, outfitting, design and technical specifications (all related) are leading to operational adjustments throughout the project delivery. Such changes are often driven by external factors like uncertainty in market demand, regulatory interventions and technological innovations. One example is shipbuilding for advanced marine operations (Emblemsvåg, 2014), with an exploratory study of a large, dynamically changing project in Hansen et al. (2020). This type of uncertainty represents a substantial challenge in an increasing number of projects (Böhle et al., 2016; Atkinson et al., 2006). It is difficult to anticipate and describe statistically, and it may lead to changes in the work content, and subsequently, to changes in requirements with alterations in the project network itself, i.e., in the activities to be performed and their sequencing (Hazir and Ulusoy, 2019; Vaagen et al., 2017). This comes on top of uncertain and usually correlated activity durations, for any fixed design. The resulting dependencies in the planning problem are, therefore, very complicated. A practical example is related to repurposing and reoutfitting ships with competing technologies with

uncertain performance, introduced in Section 5 and discussed in Section 6.

The negative impact of disturbances and time delays in design projects (Nichols, 1990) forces developers to consider managerial flexibility (Huchzermeier and Loch, 2001), through alternative implementation paths and the option to delay the choice into the project delivery (Ibadov and Kulejewski, 2019), and alternative technologies producing the same result (Creemers et al., 2015). From real option approaches to investments, we know that higher uncertainty in the payoff—for example, from design changes, usually with a defined customer value—increases the value of flexibility (Dixit et al., 1994). One implication of this insight is that the more uncertain the project payoff is, the more efforts should be made to develop flexibility to enable changing the direction of the project (Huchzermeier and Loch, 2001). But how much flexibility? Simchi-Levi (2010) shows that lower levels of flexibility may nearly capture the benefits of full flexibility. There is also evidence that higher operational uncertainty may reduce the option value, see Creemers et al. (2015) for a numerical analysis. In line with this, Vaagen et al. (2017) show that flexibility through design and process modularization is less valued when a ‘safer’ alternative is available; referring to the situation of a less standardized modular architecture.

\* Corresponding author.

\*\* Corresponding author.

E-mail addresses: [Michal.Kaut@sintef.no](mailto:Michal.Kaut@sintef.no) (M. Kaut), [Hajnalka.Vaagen@ntnu.no](mailto:Hajnalka.Vaagen@ntnu.no) (H. Vaagen), [Stein.Wallace@nhh.no](mailto:Stein.Wallace@nhh.no) (S.W. Wallace).

Another insight from real options to investments is that negative correlations provide flexibility, by hedging or switching options (e.g. switching between alternative technologies) (King and Wallace, 2012). When correlations are modelled, the planning strategy shows *not* to be sensitive to the marginal distributions (Vaagen and Wallace, 2008). This supports Wall (1997) on the potentially higher impact of correlations on project performance over distributions of task duration uncertainties. Moreover, due to potential changes in the project network we expect uncertainty in design to have higher impact on performance than uncertainty in activity durations (Zhu et al., 2005; Vaagen et al., 2017).

But what type of uncertainty is critical for prioritization of project tasks, and under which circumstances? We have limited understanding of the relationships among these types of uncertainties, which we consider as a gap in the literature. As a consequence, the relationship between operational uncertainty and the value of managerial flexibility through the option to delay decisions or switch between alternatives is less clear too. For example, we know that buffer management—commonly used to handle project uncertainty (Van de Vonder et al., 2006)—has limited value for design uncertainty (as we do not know where and how much buffer is needed), but should we respond to uncertainty in activity duration in the same way as to design changes, by e.g. postponing some decisions?

Answers to these questions would help better understand where and what type of preventive efforts to allocate. That said, most models investigate a single source of uncertainty (Hazir and Ulusoy, 2019), and the research efforts on including alternative designs, technologies and implementation paths are also limited (Servranckx and Vanhoucke, 2019). One reason may be the limited scope of traditional project management, failing to encompass all phases in the project life cycle (Atkinson et al., 2006), this leading to a disconnect between project scope, design and planning functions (Gunasekaran and Ngai, 2012).

Studies on multiple sources of uncertainties are largely limited to investigating resource availability with random activity durations (Hazir and Ulusoy, 2019). We have found two exceptions of relevance to us. One is the Özdamar and Alanya (2001) paper, dealing with activity duration and requirement uncertainty simultaneously, in the context of software development. There, both sources of uncertainties are represented as fuzzy numbers, and a simple generic heuristic algorithm based on four priority rules is proposed to prepare minimal timespan project plans. The second is Creemers et al. (2015), investigating project scheduling with stochastic activity durations while considering alternative technologies to reach the project objectives. Technology success is presented by a probability assigned to each activity making up a particular technology (not unlike the way design uncertainty is presented in Vaagen et al. (2017)). The authors show that managerial flexibility may be too costly to handle high operational uncertainty. None of those models provide a systematic decision support framework, though, to help deepen insight into the relationships among the different sources of uncertainties. Nor do they handle critical modelling aspects, like the planning dynamics driven by information arrival, and correlations, which we discuss in the following sections.

Given the above, the purpose of our paper is to provide a model that helps investigating the combined impact of, and relationships between, design changes and stochastic activity durations, including correlations between these.

To achieve our aim, we need a modelling framework that explicitly handles the two-level project uncertainty, with stochastic and correlated activity durations *conditioned* by uncertainty in design. The three main aspects to deal with simultaneously are arrival of information and future decisions, as well as dependencies (we only study correlations) between stochastic activities.

The development of a stochastic programming framework with these aspects is our *first* contribution. The *second* is insight into the relationship between uncertainty that leads to network changes (may this come from design or technology uncertainty) *and* variation in activity duration. This is achieved through numerical experiments. *Third*, we improve

understanding on the general effects of correlations on planning.

The remainder of the paper is organized as follows. The literature on modeling the sources of uncertainties under investigation in relation to the important modelling aspects are discussed in Section 2. In Section 3 we provide the justification for the choice of stochastic programming as our modelling approach. The modelling approach with the full model is described in Section 4. Section 5 is dedicated to the test cases and results. Managerial implications are given in Section 6. We conclude in Section 7.

## 2. Literature review

The distinction between uncertainty in activity duration, where activity times or resource demands may change, *and* uncertainty in requirements, where new activities or precedence relations may be added or deleted in the network, is made in Zhu et al. (2005), and later in the review paper of Hazir and Ulusoy (2019). In the latter, distinction is made between requirement uncertainty, as internal with a certain organisational ability to control it, *and* uncertainty in market-, technological- and regulatory conditions as external with limited predictability and limited ability to control it. As introduced in Section 1, in complex engineer-to-order projects these external factors are exactly those leading to potentially high impact changes in requirements through changes in design and work content, and in the project network itself, in order. See Vaagen et al. (2017) for a discussion. As such, in this paper we differentiate between (i) uncertainty in activity durations, predictable up front and statistically describable, commonly handled by buffering around critical path approaches (Van de Vonder et al., 2006), *and* (ii) potentially high impact requirement uncertainty with limited predictability, most often handled reactively after a change has been materialized (Petit and Hobbs, 2010; Hällgren and Maaninen-Olsson, 2005). This distinction is motivated by Simchi-Levi et al. (2015). While that discussion is within the context of complex dynamically changing supply chains, it applies to projects as well.

A second uncertainty classification of relevance is that between the negative and positive sides of uncertainty (Atkinson et al., 2006), and the consequence of this on plans and strategies. Chapman and Ward (2011) and Ward and Chapman (2003) argue that traditional project risk management is overly focused on the risk or threat (which is primarily cost-driver), with limited ability to capture the opportunity (which is primarily profit driver) side of uncertainty. Opportunities may arise from market demand and technological uncertainty. Regulatory interventions (such as new technology standards) are primarily downside risks, but that can sometimes be turned into opportunities (Loch et al. (2011), e.g. p. 5), such as low-emission technology that may increase customer value and may open up for new markets.

Other classifications can be found in Ward and Chapman (2003), with distinction between variability of project estimates, uncertainty around the basis for estimates, design and logistics uncertainty, and uncertainty related to the relationship between stakeholders. Pich et al. (2002) distinguish between variations as random deviations with smaller impact, foreseen uncertainty, unforeseen uncertainty, and chaos. Chaos happens rarely and has potentially high impact on project targets. Hansen et al. (2020) discuss a complex example of ‘chaos’, timely handled by team collective intelligence and lean construction practices.

On the more traditional project uncertainty classification approaches we mention Hällgren and Maaninen-Olsson (2005), with distinction between risks, changes and deviations: risks as known yet unrealized situations (managed by traditional risk management approaches), changes as realized situations with a significant divergence to the project plan (i.e., managed reactively), and deviations as situations, regardless of consequence, that deviates from any plan in the project.

Research on *requirement uncertainty and alternative (design) solutions* in planning is limited and largely based on small case examples developed independently, see for example Servranckx and Vanhoucke (2019).

They extend the resource-constrained project scheduling problem to allow for flexibility through multiple networks. Traditional resource-constrained scheduling problems assume deterministic project structure with a fixed set of activities, but there are situations when activities can be excluded from the final schedule. This is known as resource-constrained project scheduling with alternative subgraphs. The authors apply a tabu-search procedure to the selection of alternatives and the scheduling of the chosen alternative, and provide a systematic theoretical framework to the problem with multiple types of alternative subgraphs (nested and linked).

Ibadov and Kulejewski (2019) propose an alternative network model with a fuzzy decision node, to model independent and alternative activities, when the plan is expected to change in relation to the initial one. The plan update is based on choosing a predicted alternative path. This approach makes it possible to analyse multiple alternatives, in terms of the relevant characteristic of the construction project and the conditions for implementation. The computational complexity of including alternative variants is acknowledged, while also stating that information on variant preferences gathered at the decision nodes suggests which network variant to solve first.

Planning modular projects with alternative technologies is found in the stochastic programming approach of Creemers et al. (2015), with modules assumed to be independent parts, and their alternatives possibly active in parallel. Other papers observe that design processes include different types of logical relations between activities; e.g., the resource constrained project scheduling problem with logical constraints in Vanhoucke and Coelho (2016).

Vaagen et al. (2017) propose a stochastic mixed integer optimization program to study the impact of design uncertainty on project performance, under the assumption of deterministic activity durations. They show how delaying design decisions plays a role, and report the quantified impact of proactive strategies with options, with about 35% lowered expected costs as compared to reactive strategies with deterministic network plans updated in light of change. That model is developed for a principal study of small problem instances of the true complexity, and not suited for large applications.

We also mention early work on modelling design risks and uncertain number of design iterations to meet a design criteria (Luh et al., 1999) and stochastic project networks (Neumann, 1990).

*Stochastic activity durations* is the most frequently studied source of uncertainty in classical project scheduling, with PERT network models as the dominant approach (Lambrechts et al., 2010; Van de Vonder et al., 2006). A large share of this research assumes a static environment with known project structure (see a discussion in Servranckx and Vanhoucke (2019)), but project activities are often subject to substantial uncertainty, leading to schedule disruptions. These models do not handle alternative solution paths, and design uncertainty cannot be properly handled by scaling stochastic activity durations and buffering for fixed networks.

Approaches developed to handle randomness in activity durations also fail to properly handle the planning dynamics, information arrival and future decisions simultaneously. Two important research streams in this direction are proactive-reactive scheduling (Herroelen and Leus, 2005; Van de Vonder et al., 2006; Artigues et al., 2005) and stochastic resource-constrained scheduling (Herroelen et al., 2002). These approaches are dealing with a sequence of decisions from static models and are, hence, not flexible, despite the alternatives provided in contingent planning. One exception is found in Deblaere et al. (2011). This approach uses an optimized decision rule within a simulation model to estimate changes in parameter values, and achieves a near optimal setting. While it handles the dynamics of the problem, i.e., information arrival and future decisions simultaneously, and is shown to outperform many alternative approaches, it cannot say how good the optimized decision rule is. It can only compare it with others.

In general, the possibility to have future decisions conditioned on new information (e.g., changes in design or the progress of activities) is

lacking in classical project scheduling models. The major difficulty is that there is no arrival of information in these models, and no flexibility to adapt changes. Reaction to change, by rerunning deterministic models based on new information, is done (Jørgensen and Wallace, 2000). But such reactive approaches with a sequence of deterministic decisions, are not flexible and have potentially very high adaptation costs (King and Wallace, 2012; Vaagen et al., 2017).

Moreover, most scheduling models for stochastic activity durations also lack a discussion on *correlations*. Sequences driven by design and engineering constraints are commonly addressed and expressed in project networks, but dependencies of the type of correlations are less incorporated (Kadane and Wolfson, 1998). In general, correlation studies in project management are few and mainly limited to simulation (Khodakarami and Abdi, 2014), for example Monte Carlo (Chapman and Ward, 2011), but these have theoretical limitations in modelling complex cause and effect relationships. Khodakarami and Abdi (2014) propose a quantitative assessment framework that makes it possible to incorporate uncertainty and causality in project cost estimation. They integrate the inference process of a Bayesian network with the traditional probabilistic risk analysis. Based on simulation, Wall (1997) concludes with large errors when models do not consider correlations, and claims correlations to be more important than the distributions representing task duration uncertainties. The general effect of correlations is that positive correlations increase risk, while negative correlations reduce risk and are perceived as free hedging. If these are not captured, the final project duration distribution will not provide a correct understanding on the true uncertainty. From correlation studies in other fields we know that these may have very high impact on planning decisions (Vaagen and Wallace, 2008). For sources of project correlations see e.g., Schuyler (2001).

For a comprehensive, although not exhaustive, review on classification and methods for modeling project uncertainty see Hazir and Ulusoy (2019).

### 3. The choice of stochastic programming as modelling approach

As introduced earlier, we need a modelling framework that handles arrival of information and future decisions simultaneously, as well as correlations between stochastic activities. In this section we argue for the choice of stochastic programming as appropriate for this purpose.

We know that delaying design decisions into the project delivery period can add value to a project, but only if the delay allows the implementation of a better alternative without disturbing the delivery process substantially. To enable this, a planning model where decisions have the potential to be changed when new information becomes available is needed; i.e., a proactive approach that takes both arrival of information and future decisions that might unfold into account. There are very few such approaches in the literature. The simulation approach with optimized decision rules in Deblaere et al. (2011) handles information arrival and future decisions simultaneously, but we do not know how far the solution is from optimal. Vaagen et al. (2017) applies stochastic optimization to handle arrival of design choice information (but not the arrival of information about activity durations). This model handles the true complexity, but for small problem instances and conceptual studies only. Due to complexities involved, the general formulation of this stochastic dynamic problem is stated as unsolved in Jørgensen and Wallace (2000), and still it is Vaagen et al. (2017). That said, conceptual knowledge developed by small stochastic problem instances on what makes solutions good, can help finding good solutions without actually solving stochastic programs King and Wallace (2012).

Correlations further complicate the modelling problem, and are hard to deal with analytically. Simulation is therefore dominant for project correlations studies (Khodakarami and Abdi, 2014). Simulation helps to establish understanding on the project risk and on the effects of potential decisions, before the decisions are made. But the models do not provide explicit suggestions for what decisions to make, and they lack the

connection between future and present decisions; hence, not appropriate for the purpose of this paper. In the fields of stochastic network design and product- and assortment planning, King et al. (2012), Vaagen et al. (2011b) and Lium et al. (2007) point to numerical stochastic programming as the method suited to handle complicated uncertainty patterns and correlations. Lium et al. (2009) suggest that by consolidating two negatively correlated demands, flexibility and free hedging, as well as an effective use of capacities, can be achieved in network design. In situations with strong positive correlations among high-probability high-demands, flexibility shows to have less value, and the authors suggest schedules that accommodate the most probable scenarios with most demands being high at the same time, using buffers. Vaagen and Wallace (2008) formulate a product-line planning problem with bimodal distributions and correlations, and show that flexibility and hedging is mainly driven by uncertainty in design with respect to the future state in the market as preferred/not preferred (modelled by correlations), and not very sensitive to the specific values of the marginal demand distributions for a given preference. Moreover, Vaagen et al. (2011a) show that there is high value in pairing products that are negatively correlated and also substitutable to some extent in the product-line (i.e., perceived as alternatives from the customer perspective). This latter research stream in a product-line context, treats a problem with two-level uncertainty and complex dependencies. It is similar in spirit to the one at hand, with alternative designs on a higher level, and on lower level, for a given design, statistically describable demand uncertainty (in product portfolios) and activity duration uncertainty (in projects).

The current paper is founded on the above research efforts to use stochastic optimization to handle complex uncertainty patterns with correlations, and to model arrival of information and future decisions. While computationally demanding for large problems, it is appropriate for the purpose of this paper, to develop conceptual learning on the combined impact of critical sources of project uncertainties.

#### 4. Stochastic-programming formulation

In this section, we describe our stochastic-programming model. While we extend the problem in Vaagen et al. (2017) to also handle stochastic and correlated activity durations, we devise a new stochastic model and use a different way of modelling the activity progress in order to reduce the number of binary variables in the model. Before we present the model formulation itself, we describe its most challenging part, modelling of the stochastic activity durations.

##### 4.1. Modelling of stochastic activity durations

Stochastic activity durations are principally different from uncertainty usually handled in stochastic optimization models, because we have to start an activity in order to learn its duration; information arrives as a consequence of decisions, not just because time has passed. In other words, we are dealing with *endogenous*, or *decision-dependent*, uncertainty (Jonsbråten et al., 1998), specifically the Type-2 endogenous uncertainty (Goel and Grossmann, 2006). In the context of scenario trees, this would correspond to a tree where the time of the branching depends on decision variables in the model. Moreover, this endogenous uncertainty has to be combined with any exogenous uncertainty we might have; in our case, this means the stochastic design changes.

As described above, we have two types of uncertainty: ‘standard’ exogenous uncertainty related to design choices, modelled by a scenario tree, and the endogenous uncertainty of activity durations. To model this double uncertainty, we use an approach similar to the one from Goel and Grossmann (2004, 2006): the design-choice uncertainty is described by the set  $\mathcal{N}$  of scenario-tree nodes, while the stochastic durations are modelled using multiple copies of the this tree, referred to as *duration scenarios* and indexed by  $s \in \mathcal{S}$ . We then add constraints connecting the same nodes of different copies of the tree, enforcing equal decision at the

nodes as long as we have not learned the duration of the corresponding activity. These constraints are like the usual non-anticipativity constraints (NACs), except that they are being switched on and off by the decision variables; for this reason, we call them *dynamic non-anticipativity constraints* (DNACs).

This is illustrated in Fig. 1, for a case of a single activity  $a$  with two possible durations. There, we start with the standard scenario tree (Fig. 1a), describing all the exogenous uncertainty. Since  $a$  has two possible durations, we duplicate the tree, and add DNACs between all corresponding nodes (Fig. 1b). Note that the constraint connecting the root nodes is marked as active, since the first-stage decisions are unique by definition. Now, let’s assume that  $a$  is such that we learn its duration in the first period following its start, and that we start it in the node marked in Fig. 1c. In that situation, the scenario tree from Fig. 1b will take the form shown in Fig. 1d: the DNACs are active in all nodes except the two descendants of the (marked) starting node. This means that only in those two nodes are we allowed to make different decisions, based on duration of  $a$ .

Things get more complicated if we have more activities with stochastic durations. As an example, consider the case with 2 activities with two possible durations, short (S) and long (L). To model these using our approach, we need four copies of the original scenario tree, with the following combination of durations: (S,S), (S,L), (L,S), (L,L). Then we need to add DNACs for each activity, connecting trees that differ only in duration of that activity. This means two constraints for the first activity, connecting trees (1,3) and (2,4), and two for the second activity, connecting trees (1,2) and (3,4).

In the general case, presented below, this is modelled using the set  $\mathcal{E}$  of scenario pairs connected by a DNAC, and a parameter  $A_{s_1, s_2}$  representing the activity that distinguishes the scenarios  $(s_1, s_2)$ . In the example above, we have  $\mathcal{E} = \{(1, 2), (1, 3), (2, 4), (3, 4)\}$ ,  $A_{1,3} = A_{2,4} = a_1$  and  $A_{1,2} = A_{3,4} = a_2$ .

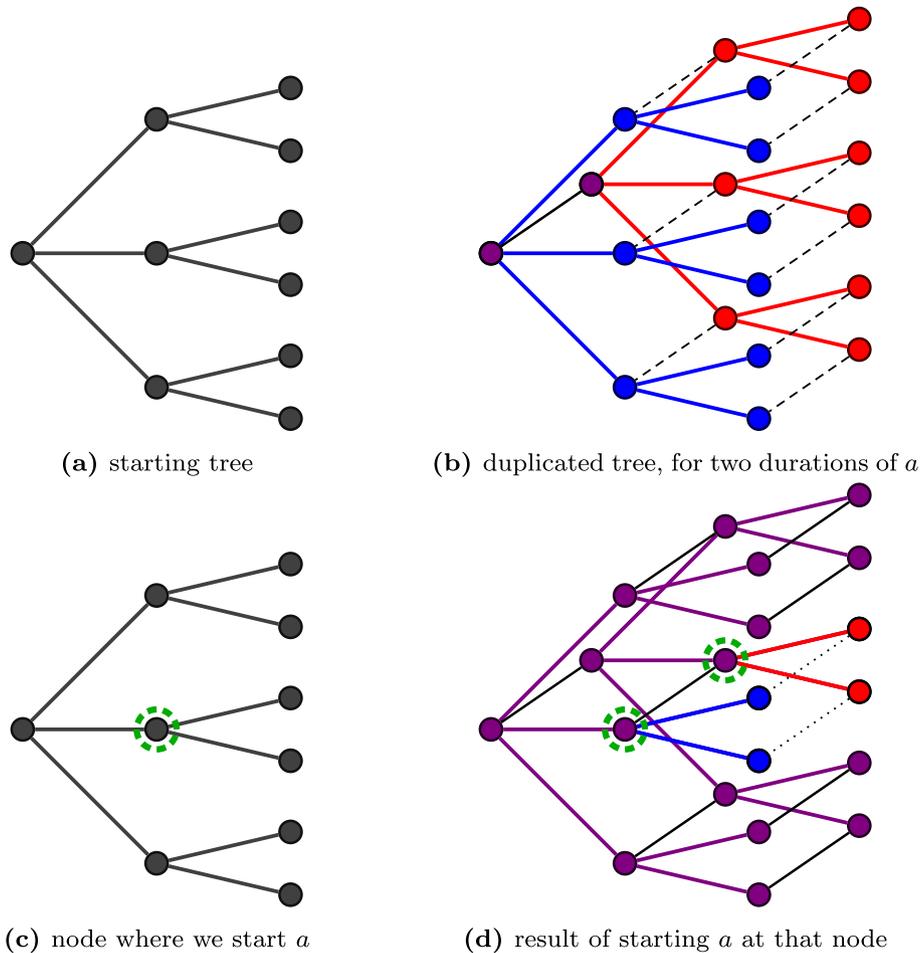
#### 4.2. Notation

##### 4.2.1. Sets

Name	Description
$\mathcal{A}$	set of all activities
$\mathcal{N}$	set of all scenario-tree nodes
$\mathcal{R}$	set of all resources
$\mathcal{A}_1 \subset \mathcal{A}$	indicator activities – no duration
$\mathcal{A}_R \subset \mathcal{A}$	real activities (with duration); $\mathcal{A}_R = \mathcal{A} \setminus \mathcal{A}_1$
$\mathcal{A}_U \subset \mathcal{A}_R$	activities that undo/reverse the results of other
$\mathcal{A}_C(a) \subset \mathcal{A}_R$	activities that conflicts with (must be undone for $a$ to start)
$\mathcal{A}_P(a) \subset \mathcal{A}_R$	activities that cannot run parallel to $a \in \mathcal{A}_R$
$\mathcal{D}_a$	set of activities that $a \in \mathcal{A}$ depends on; all must be finished
$\mathcal{D}_a^U$	set of activities that $a \in \mathcal{A}$ depends on; at least one must be finished
$\mathcal{N}_L \subset \mathcal{N}$	set of leaf nodes, i.e., nodes without children
$\mathcal{N}_l^P \subset \mathcal{N}$	set of nodes on path from the root to leaf node $l \in \mathcal{N}_L$
$\mathcal{I}_r$	a set of intervals for piecewise-linear costs of resource $r \in \mathcal{R}$
$\mathcal{D}_{a,n}^n$	stochastic variant of $\mathcal{D}_a^n$ – value at given node $n$
$\mathcal{D}_{a,n}^u$	stochastic variant of $\mathcal{D}_a^u$ – value at given node $n$
$\mathcal{D}_{a,n}^-$	stochastic dependency of type ‘not’ – conflicting activities
$\mathcal{S}$	set of scenarios for activity duration
$\mathcal{E}$	the complete stochastic event: $\mathcal{E} = \mathcal{N} \times \mathcal{S}$
$\mathcal{A}^s \subset \mathcal{A}$	activities with stochastic durations
$\mathcal{E}$	scenario pairs $(s_i, s_j)$ connected by a DNAC

As described above, we have two types of uncertainty: ‘standard’ exogenous uncertainty modelled by a scenario tree, and the endogenous uncertainty of activity durations. The former is described by the set  $\mathcal{N}$  of scenario-tree nodes, while the stochastic durations are modelled using multiple copies of the this tree, referred to as *duration scenarios* and indexed by  $s \in \mathcal{S}$ . This means that the each node of the combined tree  $\mathcal{E} = \mathcal{N} \times \mathcal{S}$  uses a double-index  $(n, s)$  for indexing.

We allow for increasing resource costs, to be able to use extra re-



**Fig. 1.** Scenario tree for a case with one activity  $a$  with two possible durations, denoted by blue and red. Dashed lines in (b) denote all the possible dynamic non-anticipativity constraints (DNACs). In (b) and (d), the full and dotted lines denote active and inactive DNACs, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

sources (tools or people), for an extra cost. This is modelled by a set of cost levels  $\mathcal{L}_r$  for each resource  $r \in \mathcal{R}$ , together with the amount  $\bar{L}_{r,l}$  of the resource available in level  $l \in \mathcal{L}_r$ , and its cost  $C_{r,l}^R$ .

**4.2.2. Parameters**

Name	Description
$P(n)$	Probability of node $n \in \mathcal{N}$ .
$n^- \Delta t$	predecessor of node $n$ , $\Delta t$ periods before node $n$
$n^-$	parent node of $n$ ; special case of $Predn\Delta t$ with $\Delta t = 1$
$t(n)$	period of node $n$
$D_n^P$	duration of period represented by node $n$
$T_0$	the first period
$A_F$	the final activity – finishing this marks the end of the project
$\bar{L}_{r,l}$	upper bound of resource $r$ in cost level $l \in \mathcal{L}_r$
$R_{a,r}$	amount of resource $r \in \mathcal{R}$ used by activity $a \in \mathcal{A}_R$ per time period
$R_{a,r}^r$	amount of resource $r \in \mathcal{R}$ used when reverting activity $a \in \mathcal{A}_R$
$A_U(a) \in \mathcal{A}$	for $a \in \mathcal{A}_U$ , this is the activity $a$ undoes/reverts
$U_a$	Multiplier for duration of undo-activities
$C_{r,l}^R$	cost of using resource $r \in \mathcal{R}$ with cost level $l \in \mathcal{L}_r$
$C_t^E$	cost of finishing the whole project at the end of $t$
$D_{a,s}^A$	duration of activity $a$ in scenario $s$
$A_{s_1,s_2}$	activity that distinguishes two scenarios $(s_1, s_2)$
$\alpha$	fraction of duration needed to distinguish between two durations

**4.2.3. Variables**

Name	Description	Range
$p_{a,n,s}$	progress of activity $a \in \mathcal{A}_R$ at the end of $(n, s)$	$\mathcal{R}_0^+$
$w_{a,n,s}$	time spend working on $a \in \mathcal{A}_R$ during period $(n, s)$	$\mathcal{R}_0^+$
$r_{a,n,s}$	time spend reverting/undoing $a \in \mathcal{A}_R$ during period $(n, s)$	$\mathcal{R}_0^+$
$f_{a,n,s}$	has activity $a \in \mathcal{A}$ finished by the end of period $(n, s)$ ?	$\{0, 1\}$
$g_{a,n,s}$	is activity $a \in \mathcal{A}$ ongoing (has non-zero progress) in per. $(n, s)$ ?	$\{0, 1\}$
$u_{r,l,n,s}$	amount of resource $r$ at cost-level $l$ used in per. $(n, s)$	$\mathcal{R}_0^+$
$c_{a,n,s}$	cumulative progress of $a$ at the end of $(n, s)$	$\mathcal{R}_0^+$
$d_{a,n,s}$	indicator for DNAC constraints for $a \in \mathcal{A}^s$ during $(n, s)$	$\{0, 1\}$

Unlike the model from Vaagen et al. (2017), which tracks the activity using binary indicators, we model the progress of each activity as a continuous variable. Even if these have to be connected to binary indicators to model dependencies etc, the new model has fewer binary variables and is therefore easier to solve. Note that it is also more flexible, as it allows for working on an activity for only a fraction of a time period, which was not possible with the previous model. However, the dependencies are only resolved at the period boundaries: if activity  $a$  finishes in the middle of period  $t$  and activity  $b$  depends on  $a$ , then  $b$  will be allowed to begin first at the start of period  $t + 1$ .

The last two variables (cumulative progress  $c_{a,n,s}$  and DNAC-indicator  $d_{a,n,s}$ ) are used for modelling of the stochastic durations, as described below.

### 4.3. The model

This section presents the objective function and constraints of the model. Throughout the section, we simplify the notation by assuming that expressions with non-existing values (such as  $n^-$  in the root node) evaluate to zero.

#### 4.3.1. Objective function

$$\text{minimize } \sum_{(n,s) \in \mathcal{E}} P(n) \left[ \sum_{r \in \mathcal{R}} C_{r,l}^R u_{r,l,n,s} + C_{l(n)}^E (f_{A_{F,n,s}} - f_{A_{F,n^-,s}}) \right] \quad (1)$$

The objective is to minimize the expected costs, consisting of the resource-usage costs and extra penalty term depending on the finishing time of the whole project, i.e., end time of the final activity  $A_F$ .

#### 4.3.2. Activity progress constraints

$$p_{a,n,s} = p_{a,n^-,s} + w_{a,n,s} - 1 / U_a \quad r_{a,n,s} \quad a \in \mathcal{A}_R, (n,s) \in \mathcal{E} \quad (2)$$

$$c_{a,n,s} = c_{a,n^-,s} + w_{a,n,s} \quad a \in \mathcal{A}_R, (n,s) \in \mathcal{E} \quad (3)$$

$$f_{a,n,s} \leq p_{a,n,s} / D_{a,s}^A \quad a \in \mathcal{A}_R, (n,s) \in \mathcal{E} \quad (4)$$

$$g_{a,n,s} \geq p_{a,n,s} / D_{a,s}^A \quad a \in \mathcal{A}_R, (n,s) \in \mathcal{E} \quad (5)$$

$$f_{a,n,s} \geq f_{a,n^-,s} \quad a \in \mathcal{A}_R, (n,s) \in \mathcal{E} \quad (6)$$

$$f_{A_F,l,s} = 1 \quad l \in \mathcal{N}_L, s \in \mathcal{S} \quad (7)$$

Constraints (2) and (3) model the normal and cumulative progress, respectively, while (4) and (5) define indicators for finished and ongoing projects, respectively. Constraints (6) ensure that an activity marked as finished will remain so to the end. Without this, an activity could be marked as finished (and therefore could trigger start of another activity) and then reverted. Finally, (7) ensures that the project finishes in all scenarios.

#### 4.3.3. Dependencies and conflicts between activities

$$w_{a,n,s} \leq D_n^P f_{a',n^-,s} \quad a \in \mathcal{A}_R, (n,s) \in \mathcal{E}, a' \in \mathcal{D}_a^U \cup \mathcal{D}_{a,n}^U \quad (8)$$

$$w_{a,n,s} \leq D_n^P \sum_{a' \in \mathcal{D}_a^U \cup \mathcal{D}_{a,n}^U} f_{a',n^-,s} \quad a \in \mathcal{A}_R : \left| \mathcal{D}_a^U \cup \mathcal{D}_{a,n}^U \right| > 0, (n,s) \in \mathcal{E} \quad (9)$$

$$w_{a,n,s} \leq D_n^P (1 - g_{a',n,s}) \quad a \in \mathcal{A}_R, a' \in \mathcal{A}_C(a), (n,s) \in \mathcal{E} \quad (10)$$

$$f_{a,n,s} \leq f_{a',n^-,s} \quad a \in \mathcal{A}_1, (n,s) \in \mathcal{E}, a' \in \mathcal{D}_a^U \cup \mathcal{D}_{a,n}^U \quad (11)$$

$$f_{a,n,s} \leq \sum_{a' \in \mathcal{D}_a^U \cup \mathcal{D}_{a,n}^U} f_{a',n^-,s} \quad a \in \mathcal{A}_1 : \left| \mathcal{D}_a^U \cup \mathcal{D}_{a,n}^U \right| > 0, (n,s) \in \mathcal{E} \quad (12)$$

$$f_{a,n,s} \leq 1 - g_{a',n,s} \quad a \in \mathcal{A}_1, a' \in \mathcal{D}_{a,n}^-, (n,s) \in \mathcal{E} \quad (13)$$

$$w_{a,n,s} + r_{a,n,s} + w_{b,n,s} + r_{b,n,s} \leq D_n^P \quad a \in \mathcal{A}_R, b \in \mathcal{A}_P(a), (n,s) \in \mathcal{E} \quad (14)$$

$$w_{a,n,s} = 0 \quad a \in \mathcal{A}_R, (n,s) \in \mathcal{E} : \sum_{l \in \mathcal{N}_L, n \in \mathcal{N}_P^l} P(l) \left| \mathcal{D}_{a,n}^U \ominus \mathcal{D}_{a,l}^U \right| > 0 \quad (15)$$

$$f_{a,n,s} = 0 \quad a \in \mathcal{A}_1, (n,s) \in \mathcal{E} : \sum_{l \in \mathcal{N}_L, n \in \mathcal{N}_P^l} P(l) \left| \mathcal{D}_{a,n}^U \ominus \mathcal{D}_{a,l}^U \right| > 0 \quad (16)$$

Constraints (8) and (9) model the 'and'- and 'or'-type dependencies for real activities, i.e., cases where one activity depends on either all, or at least one, of a specified set of activities. Constraints (10) model conflicting activities (which can be viewed as 'not'-type dependencies),

where we cannot work on an activity as long as another one is ongoing. Constraints (11)–(13) do the same for indicator activities; since these do not have a duration, the constraints work directly on the activity-finish indicators. In addition, constraints (14) model the case where some activities are forbidden to run in parallel (at the same time).

Finally, constraints (15) and (16) ensure that an activity with stochastic dependencies cannot start before the relevant uncertainty is revealed; (15) is for real activities and (16) is for indicators. Inside the sum, ' $\ominus$ ' denotes the symmetric difference of two sets, so  $\left| \mathcal{D}_{a,n}^U \ominus \mathcal{D}_{a,l}^U \right| > 0$  is equivalent to  $\mathcal{D}_{a,n}^U \neq \mathcal{D}_{a,l}^U$ . In other words, we allow positive progress of activity  $a$  in node  $n$  only if the dependency does not change after node  $n$ .

#### 4.3.4. Resource usage

$$u_{r,l,n,s} \leq \bar{L}_{r,l} \quad r \in \mathcal{R}, l \in \mathcal{L}_r, (n,s) \in \mathcal{E} \quad (17)$$

$$\sum_{a \in \mathcal{A}_R} \left( R_{a,r} w_{a,n,s} + R_{a,r}^r r_{a,n,s} \right) = \sum_{l \in \mathcal{L}_r} u_{r,l,n,s} \quad a \in \mathcal{A}_R, (n,s) \in \mathcal{E} \quad (18)$$

These constraints track the resource usage of all activities. Together with upper bounds on  $u_{r,l,n,s}$  they ensure that we only use the resources we have.

#### 4.3.5. Decision-dependent non-anticipativity constraints

$$d_{n,s_1,s_2} \leq \frac{c_{a,Prev_{n,s_1}}}{\alpha \min(D_{a,s_1}^A, D_{a,s_2}^A)} \quad n \in \mathcal{N}, (s_1, s_2) \in \mathcal{E}, a = A_{s_1,s_2} \quad (19)$$

The remaining constraints are all for  $n \in \mathcal{N}$  and  $(s_1, s_2) \in \mathcal{E}$ , in addition to the specified ranges:

$$p_{a,n,s_2} - p_{a,n,s_1} \leq \max\{D_{a,s_1}^A, D_{a,s_2}^A\} d_{n,s_1,s_2} \quad a \in \mathcal{A}_R \quad (20)$$

$$p_{a,n,s_1} - p_{a,n,s_2} \leq \max\{D_{a,s_1}^A, D_{a,s_2}^A\} d_{n,s_1,s_2} \quad a \in \mathcal{A}_R \quad (21)$$

$$w_{a,n,s_2} - w_{a,n,s_1} \leq D_n^P d_{n,s_1,s_2} \quad a \in \mathcal{A}_R \quad (22)$$

$$w_{a,n,s_1} - w_{a,n,s_2} \leq D_n^P d_{n,s_1,s_2} \quad a \in \mathcal{A}_R \quad (23)$$

$$r_{a,n,s_2} - r_{a,n,s_1} \leq D_n^P d_{n,s_1,s_2} \quad a \in \mathcal{A}_R \quad (24)$$

$$r_{a,n,s_1} - r_{a,n,s_2} \leq D_n^P d_{n,s_1,s_2} \quad a \in \mathcal{A}_R \quad (25)$$

$$f_{a,n,s_2} - f_{a,n,s_1} \leq d_{n,s_1,s_2} \quad a \in \mathcal{A} \quad (26)$$

$$f_{a,n,s_1} - f_{a,n,s_2} \leq d_{n,s_1,s_2} \quad a \in \mathcal{A} \quad (27)$$

Constraints (19) define indicators  $d_{n,s_1,s_2}$  for the dynamic non-anticipativity constraints (DNACs):  $d_{n,s_1,s_2} = 1$  means that we can distinguish between  $s_1$  and  $s_2$  at node  $n$ . The constraints ensure that this happens only if activities connecting scenarios  $s_1$  and  $s_2$  have reached at least fraction  $\alpha$  of the shorter duration. For example, if  $\alpha = 0.5$ ,  $D_{a,s_1}^A = 4$  and  $D_{a,s_2}^A = 6$ , we have to run the project for  $0.5 \times 4 = 2$  periods before we learn the duration and hence can distinguish the two scenarios  $s_1$  and  $s_2$ .

Finally, constraints (20)–(27) enforce the dynamic non-anticipativity, pairwise for each set of variables. For example, (20)–(21) ensure that  $p_{a,n,s_2}$  can differ from  $p_{a,n,s_1}$  only if  $d_{n,s_1,s_2} = 1$ .

## 5. Test cases

### 5.1. Motivating example

The example for test cases is an extended version of that in [Vaagen](#)

et al. (2017), where we add stochasticity and correlations to activity durations, in addition to the uncertainty in design.

Consider the example of re-outfitting a vessel with competing engine technologies with uncertain performance, electric A or hybrid-electric B. It is acknowledged that in such projects one of the technologies turns out to be more compatible with the existing solution than the other, but this understanding becomes available only after re-opening the vessel. Important outfitting decisions are hence made in light of uncertainty, and stochastic activity durations connected to the competing technologies are negatively correlated. We may also know that if one task connected to one technology (A for instance) takes longer than expected, others connected to that technology may also take longer (common cause, driving positive correlations). This practical example is discussed in relation to the findings in Section 6.

Assume the project consists of the choice of design alternative, and scheduling with three activities P, D and E; depicted in Fig. 2a. There, we have introduced an indicator activity F depending on the three activities, P, D and K. The diamond shape shows that the dependency is of type ‘and’, i.e., the activity needs all its predecessors to finish in order to start.

Assume activities P and D depend on the design choice, and that the customer decision on preferred alternative can be delayed or changed during the project duration. I.e., the choice between technologies A and B is a stochastic parameter in the model.

This gives the network in Fig. 2b, where we have introduced three new indicator activities A, B, and S. The latter is of a special type, since S depends on either A or B, dependent on the customer choice. For example, if design A is preferred, this would translate into activity S depending on activity A and hence on activities PA and DA.

Further, assume that the design-dependent activities P and D can be run in two substitutable alternatives: specialized from the start (call it ‘one-step’ version, or integral design), or modularized with a common part and a specialized part for the two designs A and B. For the latter, we can start with the standardised common part and postpone the specialisation<sup>1</sup> (call it flexible two-step version, modular or set-based engine design).

For project P, this means replacing nodes PA and PB from Fig. 2b by the network of nodes depicted at Fig. 2c. There, activities PA and PB become indicator activities with dependencies of type ‘or’, i.e., they can start when at least one of their predecessors has finished. Activity D is enhanced in the same fashion. The result is a dependency graph used in the actual test cases, see Figs. 5 and 6.

Finally, we add the uncertain durations (two possible values) for four selected activities, as follows.

First, we have the situation with uncertainty on the specialisation tasks of the alternative technologies A and B. This means that we have stochastic and correlated second-step tasks of the modular versions of activities P and D; i.e., stochastic and correlated P2A, P2B, D2A and D2B, as shown on Fig. 5.

Second, we have the situation with uncertainty only on technology B, but on both implementation alternatives. Hence, we have activities POB, P2B, D0B and D2B stochastic and correlated, as shown by Fig. 6. Recall that the first-step activities of the modular solutions, P1 and D1, are made standard for both designs A and B, and are hence deterministic. The activities making up design A are also deterministic. For a practical illustration in shipbuilding, one situation with major randomness only on activities of one of the design variants, is observed for sister vessels. By completing a first vessel, shipbuilders develop knowledge and eliminate uncertainty on the preferred design with preferred implementation path. Design changes on the second (sister) vessel (driven by e.g., the market and regulatory interventions) generate stochasticity and correlations on the new activities.

For comparability of the results with those presented in Vaagen et al.

(2017), the stochastic activity durations are built around two data sets provided in that paper, as presented in Table 1. The second data set reflects higher reactivity to change, by shorter durations on the second stage specialisation tasks, and correspondingly longer on the first stage standardised tasks. Also note that in both versions, the two-step implementation alternatives of PA, PB, DA, and DB take one period longer than the integral one-step paths.

The planning horizon consists of 11 half-week periods, so the maximal duration is 5.5 weeks. We have only one resource  $r$  and each real activity uses one unit of the resource per period. We can use up to four units of the resource in each period, where the first two units cost 1.0, the third unit 1.5, and the fourth 2.0. Since we want the project to finish as soon as possible, we use an increasing penalty for the overall finish time.

In addition, we have to model the design uncertainty. We assume that the customer prefers design alternative A, but can change the preference to either B or ‘both A and B’ during the duration of the project. We allow the change after one, two, and three weeks, i.e., after periods 2, 4, and 6. In addition, we study the effect of adding an extra week (two periods) to the most challenging scenarios. The resulting scenario trees are presented in Figs. 3 and 4.

We have run the test with an increasing probability of changing design from A to B: 1%, 5%, 10% and 20% at each branching. This means that the probability of no change decreases from 97% to 85.7%, 72.9%, and finally 51.2% in the last case. I.e., we cover the range of low to nearly full uncertainty in design preference.

### 5.2. Test 1: uncertainty on the modular implementation path of both design variants A and B

Here we assume the one-step integral (i.e., non-flexible) paths to design alternatives A and B as deterministic. We have stochasticity on the flexible two-step paths to both designs; i.e., on activities P2A, P2B, D2A, D2B (as shown on Fig. 5), each with 2 values, resulting in  $2^4 = 16$  duration scenarios. Unless specified otherwise, the two values have probability 50% for all activities.

We specify two correlation values:

- Correlations between alternative versions of the same activity: P2A vs. P2B and D2A vs. D2B;
- Correlations between two activities within one design: P2A vs. D2A and P2B vs. D2B.

The remaining correlations are fixed to zero. The one-step solutions of designs A and B, as well as the first-steps in the two-step versions, have deterministic durations. The six cases for Test 1 analysis are given in Appendix A.

### 5.3. Test 2: uncertainty on design variant B

In these tests we have uncertainty on the substitutable alternative paths to design B, i.e., on the integral (non-flexible) and modular (flexible) solutions, and pairwise correlation between these alternatives, as presented by Fig. 6.

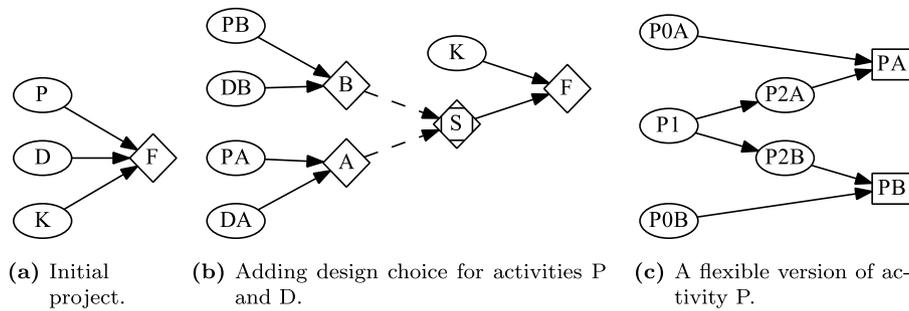
We have 4 stochastic activities POB, P2B, D0B, D2B, each with 2 values, resulting in  $2^4 = 16$  duration scenarios. Unless specified otherwise, the two values have probability 50% for all activities.

We specify two correlation values, as follows:

- Correlations between substitutable alternatives of the same activity: POB vs. P2B and D0B vs. D2B;
- Correlations between the corresponding alternatives, non-flexible and flexible, of the two activities: POB vs. D0B and P2B vs. D2B.

The remaining correlations are fixed to zero. The two cases developed for Test 2 analysis are given in Appendix A.

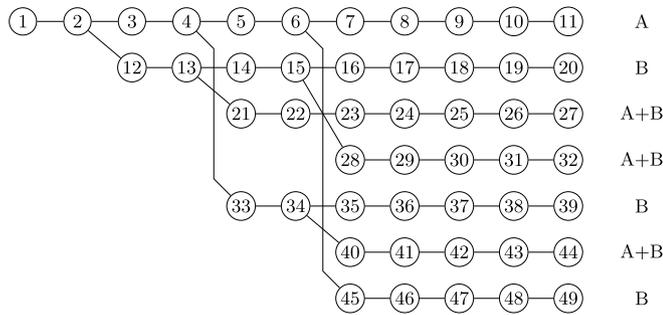
<sup>1</sup> In most cases, this extra flexibility will come at a cost.



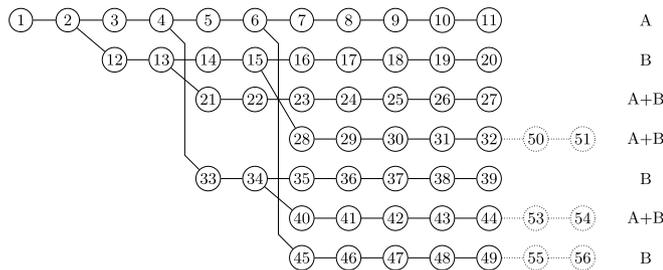
**Fig. 2.** Step-by-step construction of the motivating example. Real activities are depicted by ellipses, indicator activities with *and*-dependency by diamonds, or-dependency by rectangles, and the stochastic dependency by a combination of the two.

**Table 1**  
Deterministic activity durations from Vaagen et al., 2017.

Activity	P0A	P0B	P1	P2A	P2B	D0A	D0B	D1	D2A	D2B	K
version 1	4	3	2	3	2	3	4	2	2	3	2
version 2			3	2	1			3	1	2	



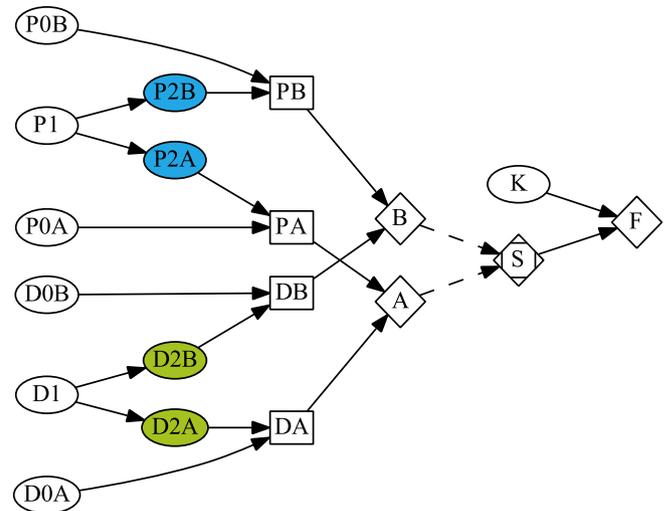
**Fig. 3.** Scenario tree 11 periods.



**Fig. 4.** Scenario tree 11 + 2 periods.

5.4. Test results

Detailed test results for the eight cases are given in 5.4. Below we summarise the results by stating that when the higher level design uncertainty is described by multiple design alternatives (A and B in our case) and delayed decisions on the final choice, randomness in activity duration and correlations between activities of alternative designs have limited impact on the planning decisions and performance. The planning guidelines suggested in Vaagen et al. (2017) to handle design change under the assumption of deterministic activity durations, are shown to be valid also under stochastic activity durations (with a few exceptions which we discuss later): Postponement is preferred whenever possible, followed by the design implementation strategy (flexible or non-flexible) that enables minimal time and costs. Flexible two-step task solutions are preferred under the possibility of quick customisation to real-time customer preferences (i.e., when the second step of a flexible



**Fig. 5.** The project network for the test with stochastic second steps of flexible paths to design A and B. Design dependent stochastic piping activities (P2A, P2B) and electro activities (D2A, D2B) are denoted in blue and green, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

solution is short relative to the first step), and in situations when extra time periods cannot be added. Non-flexible one-step task solutions combined with impact-based prioritization are preferred when there is low reactivity to real-time customer preferences (i.e., with long second steps relative to the first steps of the flexible task solutions), and when extra periods can be added to the difficult scenarios.

Exception to the above results is found in situations with correlations between the substitutable implementation paths (flexible and non-flexible) of a particular design (i.e., in Test 2 cases). In these situations, it is suggested to start implementing parts of the correlated activities to learn which one of these will have shortest completion time, before the decision on the strategy that minimizes time and costs is taken. Whenever possible, postponement is observed before learning. This learning behaviour is seen only when there is low reactivity to design change; i.e., in case 7 (see Appendix A) with long second-step durations compared to the first-step. Learning is most prominent in situations when extra time periods cannot be added to the project

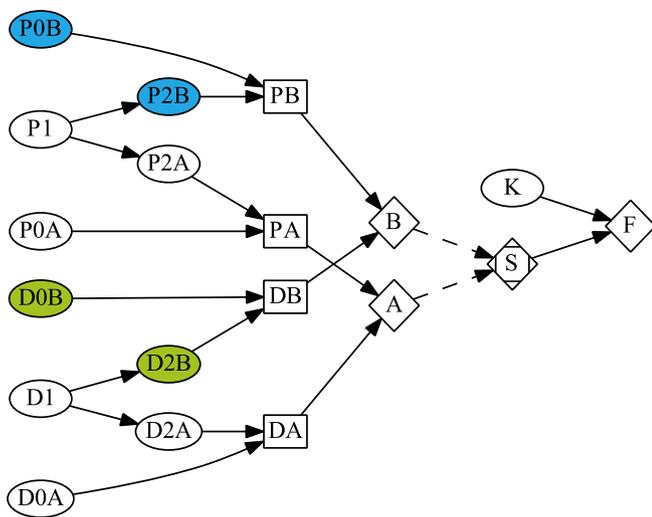


Fig. 6. The project network with stochastic design B. Stochastic piping activities (P0B, P2B) and electro activities (D0B, D2B) are denoted in blue and green, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

delivery, and is highest under high values of negative correlations.

## 6. Managerial implications and guidelines for planning decision making under uncertainty

In engineering projects, changing market conditions and uncertainty in the future performance of a design or technology are leading to frequent changes in design and technical specifications. This type of uncertainty is difficult to estimate, and is most often not addressed in advance but handled reactively after a change has materialized (Petit and Hobbs, 2010). Randomness in activity duration, on the other hand, is usually addressed early and handled by buffering around the critical path. For changes in design, buffering is suboptimal, as we don't know where to put buffers and how much. One way to handle this type of change is to consider multiple alternatives in product and process design. In flexible design strategies, like set-based or modular design, multiple alternatives are described, but it is less clear how to implement this flexibly in planning tools. Moreover, while simulation studies point to correlations as more impactful than the distribution of activity durations, planning approaches with correlations are few. Following, there is limited understanding on which one of these sources of uncertainties is more important and why, and under which circumstances. As a result, project managers may not adequately prepare for them, and preventive efforts (such as buffers or flexibility strategies) may be dysfunctionally designed and allocated. Our paper attempts to improve understanding on these managerial aspects in the following ways.

*Firstly*, the results suggest that the most critical uncertainty in planning engineer-to-order projects is the higher level design uncertainty (which we perceive as requirement uncertainty), dictating what flexibility to develop and under which circumstances. We have implemented flexible design strategies in our decision model, and the analyses show that when models are extended this way, randomness in activity durations and correlations between activities of alternative designs have limited impact. The planning guidelines suggest postponement to be followed by flexible implementation versions (we called it two-step design), whenever quick specialisation can be achieved by a short second step relative to the first step (which is what flexible design strategies strive for). When flexible task solutions are not available, managerial flexibility achieved by two-step design (meaning modularized solution) is less valued. In these situations, postponement is to be followed by impact-based prioritization.

*Secondly*, correlations between substitutable solutions for a given design, drive a certain front-end learning behaviour, as follows: First, start to implement parts of the correlated activities to learn which one of the correlated and substitutable task solutions will have shortest completion time, before the decision on the flexibility strategy that minimizes time and costs is taken. Thereafter, the implemented activities that enabled learning will be potentially, but not necessarily, uninstalled. Postponement is used before learning, whenever possible.

Learning is more prominent in situations when time is critical and extra periods cannot be added to the time horizon, and when there is low reactivity to real-time customer preferences (i.e., long second steps relative to the first steps of the flexible task solutions). The value of learning is highest under high negative correlations, i.e., when the durations of alternatives are expected to go in opposite direction.

One practical situation when starting something to learn makes sense, is when re-outfitting existing ships with competing technologies, e.g., diesel electric or diesel electric-hybrid propulsion systems. Hybrid technology provides high-efficiency alternative applications in some cases, by storing electrical energy in rechargeable batteries, but its compatibility with existing solutions on one-of-a-kind specialized projects is an acknowledged challenge. It turns out that one solution alternative (e.g., fully customized or modular) is better, with shorter expected duration than others (i.e., negatively correlated), although substitutable. This understanding only becomes available after collecting information, in our case after re-opening the ship, often far into the re-build process with a chosen solution. We may also know that if a task connected to one alternative (e.g., fully customized electro) takes longer than expected, other tasks connected to that alternative (e.g., fully customized piping) may also take longer (common cause, driving positive correlations). Information is revealed because of decisions, not just because time has passed. This means that when a particular solution alternative is fixed early (as it is in common practice), rework loops are nearly unavoidable on one-of-a-kind re-builds.

Fig. 7 illustrates a modular solution of Rolls Royce propulsion system. This enables quick adaptation from diesel-electric (a) to multiple variants of diesel electric-hybrid systems (b). The modular solution is designed to meet the compatibility challenges involved in re-outfitting existing ships and challenges imposed by late design changes in new-build ships.

In conclusion, we suggest to use knowledge of correlations to learn early the duration of uncertain activities, and postpone the selection of a final alternative (in our case, modular or fully customized) until after this learning. Conceptually, this learning behaviour is in line with decision making under uncertainty, in that it suggests reducing or eliminating uncertainty as early as possible. In our case this involves starting stochastic and correlated activities to learn. This may come with some extra early costs (as activities may be uninstalled), but enables developing solutions with reduced reworks, hence reduced time and costs. In common re-outfitting practice, the decision on a possible solution is taken before an activity is started. This leads, eventually, to learning the compatibility of the chosen solution, but this type of learning may come late and with unreasonably high costs.

Knowledge on the influence of correlations is particularly important in markets where the competitive advantage is built on handling changes throughout the project delivery, and where full and highly detailed information is not available before the building process starts.

In our model, the choice of which activity to start first, in order to maximize the value of learning, is driven by optimality criteria. The activities come from a set of activities with specified correlations. In our modelling framework, we may see that starting an activity is done just to learn. This situation can in most real cases better be represented by an actual activity with the specific purpose of learning. The handling of the stochastics would otherwise be as we already outlined. Since 'investigating' in order to learn might be simpler than starting in order to learn, the use of a learning activity is probably better. However, not to make our model even more complicated, we have chosen not to include

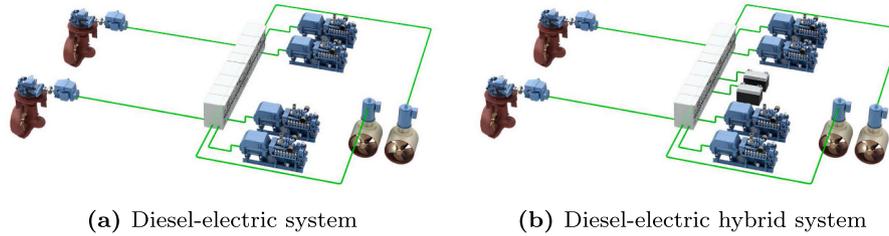


Fig. 7. Rolls Royce propulsion systems. Figures with permission from Kongsberg Maritime, © 2018 Kongsberg Maritime.

learning activities explicitly. That said, whenever the model shows a solution where learning, in fact, can be seen to be the purpose, it may make sense to actually carry out a learning activity, and not pretend to start the actual activity.

7. Conclusions

In this paper we treated decision making in project planning with uncertainty at two levels, with stochastic and correlated activity durations conditioned by the higher level design uncertainty, which we approached by stochastic dynamic programming. The model we developed is simultaneously treating information arrival and future decisions, as well as correlations between stochastic activities. Our main aim was to better understand how the different uncertainty elements, together, affect the planning process and strategies, and to understand how to use real options to switch among design alternatives, to operationalize modular or set-based (flexible) design development. The contribution of the paper is, therefore, placed within the wider context of management of project uncertainty.

While the stochastic program we provide is not for large applications (as that would be very difficult with the complex uncertainty and dependency patterns at hand), it enables conceptual learning through small problem instances of the complexity, and exemplifies how to operationalize design alternatives in project planning. The learning is

threefold: 1)insight into the impact hierarchy of the different sources of uncertainty, 2)learning about the influence of correlations and into the decision behaviour driven by these, and 3) increased insight into the structure of planning solutions under multiple sources of uncertainties. We anticipate these to be valuable for project management approaches suited for large applications, such as advanced simulation, and less tangible project management processes associated with organisational learning.

On a final note, we believe our findings *together* with earlier results on correlations in assortment planning (discussed in Section 3), contribute to improved knowledge on the general effects of correlations. Concretely, from assortment planning we know that pairing negatively correlated items that are also substitutable (to some extent) reduces the monetary portfolio risk. Our findings point to similar insight within project planning decision making, i.e., preparing with negatively correlated and substitutable alternatives of uncertain activities, can reduce project risk.

Acknowledgements

The authors thank anonymous reviewers for valuable comments. Wallace is partly supported by Grant 280536 from the Research Council of Norway.

Appendix A. Test cases

The following six test cases are developed for Test 1 analysis:

Case 1 uses the two data versions from the deterministic case on Table 1 as the two scenarios. Here, the flexible two-step versions of tasks P and D are made shorter in expectation, compared to the deterministic case; a less realistic situation.

activity	POA	POB	P1	P2A	P2B	DOA	DOB	D1	D2A	D2B
value 1	4	3	2	3	2	3	4	2	2	3
value 2				2	1				1	2
mean	4	3	2	2.5	1.5	3	4	2	1.5	2.5

Case 2 uses the first data version from Table 1, with stochastic durations equal to the mean ±1. That is, long second-steps are made uncertain.

activity	POA	POB	P1	P2A	P2B	DOA	DOB	D1	D2A	D2B
value 1	4	3	2	2	1	3	4	2	1	2
value 2				4	3				3	4
mean	4	3	2	3	2	3	4	2	2	3

Case 3 uses the same values as Case 2, but the longer variants get probability 2/3, resulting in longer expected duration of the flexible variants.

Case 4 uses the same values as Case 2, but the longer variants get probability 1/3, resulting in shorter expected duration of the flexible variants.

With respect to the second data version in Table 1, we observe that we cannot use the same logic as for the first data variant (with stochastic duration equal to the mean ±1), since we would end up with duration equal to zero. Hence, we change the definition to be the mean ± 50%. This allows to formulate the following two new cases:

Case 5 is like Case 2, but using the new stochastic durations with mean ±50% (i.e., higher uncertainty), resulting in long and highly uncertain

second-steps.

activity	P0A	P0B	P1	P2A	P2B	D0A	D0B	D1	D2A	D2B
value 1	4	3	2	1.5	1	3	4	2	1	1.5
value 2				4.5	3				3	4.5
mean	4	3	2	3	2	3	4	2	2	3

Case 6 is analogous to Case 5, but uses the second data version (with short second stages) from Table 1:

activity	P0A	P0B	P1	P2A	P2B	D0A	D0B	D1	D2A	D2B
value 1	4	3	3	1	0.5	3	4	3	0.5	1
value 2				3	1.5				1.5	3
mean	4	3	3	2	1	3	4	3	1	2

The following two cases are developed for Test 2 analysis.

Case 7 Activity durations are based on the first data variant from Table 1 (long second steps of flexible versions, i.e., low reactivity to changes), with stochastic durations equal to the mean  $\pm 50\%$ .

Case 8 is like Case 7, but based on the second data variant from Table 1 (short second steps, i.e., high reactivity to changes).

## Appendix B. Results

### Appendix B.1. Test 1, Cases 1-6

In these cases we have randomness on both designs A and B, with stochastic and correlated second-step durations of the flexible alternatives of activities P (P2A, P2B) and D (D2A, D2B), denoted in blue and green on Fig. 5.

In Case 1 the flexible two-step versions of tasks P and D are used, but this is because we have made the two-step solutions shorter (on average), compared to the deterministic one-step versions.

In Case 2, the findings support results from the model with deterministic task durations in Vaagen et al. (2017). Flexible two-step versions are less preferred when there is low reactivity to real-time customer preferences (i.e., long second step specialisation tasks). Postponement is used before impact-based prioritization of one-step task solutions. We do not see any clear impact of the correlations.

In Case 3 flexibility is even less attractive as the flexible versions are now made even longer in expectation, as compared to Case 2.

Case 4 is similar to Case 1, where flexible task solutions are used because we have made the stochastic second steps shorter in expectation.

In Case 5 the findings are in line with the Case 2 findings. Postponement is preferred before impact-based prioritization of one-step solutions. The benefits of design flexibility are even less attractive when there is high uncertainty connected to the long second-step specialisation tasks.

In Case 6 flexible two-step versions are preferred when the second step specialisation tasks are short, relative to the first steps of the flexible task solutions. Postponement and flexibility is observed in all tested data variants.

### Appendix B.2. Test 2, Cases 7-8

In Test 2 cases we only have randomness on design B, with stochastic and correlated durations of tasks P0B, P2B, D0B, D2B, denoted in blue and green on Fig. 6.

In Case 7, with long second steps of the flexible task solutions of design B (i.e., low reactivity), in situations where we allow 20% increase in time (2 periods) for the most challenging scenarios and for data variants with zero correlations, postponement is followed by nonflexible (one-step) task solutions. In tests without extra periods, we observe increased use of flexible solutions. This confirms previous findings in Vaagen et al. (2017) and findings from Section Appendix B.1.

Under data variants with correlations, a certain learning behaviour is observed before the optimal strategy that minimizes duration and costs (as a combination of flexible and non-flexible task solutions) is taken. Learning is enabled by starting the one-step versions D0B and P0B until we learn their durations. From this and knowledge on correlations we learn the duration of correlated activities, before actually starting them. For example, from learning the duration of P0B and negative correlations between P0B and P2B it then follows that if P0B is long, P2B will probably be short. A second example, from learning D0B, knowledge on positive correlations between D0B and P0B, and negative correlations between P0B and P2B, it then follows that if D0B is long, P2B will probably be short. This learning cannot be achieved if starting with the flexible two-step version of P, as in that case we have had to start with the first-step common tasks P1 and D1, which are deterministic and not correlated with P0B and D0B. After learning, the optimal strategy that minimizes duration and costs is chosen. This learning behaviour is most seen in situations when extra periods are not allowed to be added. Postponement is used when possible.

In Case 8 we have high reactivity to design change (through short second step customisation tasks of the flexible solutions), and more uncertain one-step non-flexible versions of design B. For all data variants with and without correlations, flexible two-step task solutions are preferred. The learning behaviour described for Case 7 is not detected here, and we also see less postponement. In line with our earlier findings, high reactivity to changes through short customisation tasks drives flexibility.

## References

Artigues, C., Billaut, J.-C., Esswein, C., 2005. Maximization of solution flexibility for robust shop scheduling. *Eur. J. Oper. Res.* 165 (2), 314–328.

Atkinson, R., Crawford, L., Ward, S., 2006. Fundamental uncertainties in projects and the scope of project management. *Int. J. Proj. Manag.* 24 (8), 687–698.  
 Böhle, F., Heiding, E., Schoper, Y., 2016. A new orientation to deal with uncertainty in projects. *Int. J. Proj. Manag.* 34 (7), 1384–1392.

- Chapman, C., Ward, S., 2011. How to Manage Project Opportunity and Risk. Why Uncertainty Management Can Be A Much Better Approach than Risk Management, third ed. John Wiley & Sons Ltd, West Sussex, p. 520.
- Creemers, S., De Reyck, B., Leus, R., 2015. Project planning with alternative technologies in uncertain environments. *Eur. J. Oper. Res.* 242 (2), 465–476.
- Deblaere, F., Demeulemeester, E., Herroelen, W., 2011. Proactive policies for the stochastic resource-constrained project scheduling problem. *Eur. J. Oper. Res.* 214 (2), 308–316.
- Dixit, A.K., Dixit, R.K., Pindyck, R.S., 1994. Investment under Uncertainty. Princeton university press.
- Emblemsvåg, J., 2014. Lean project planning in shipbuilding. *J. Ship Prod. Des.* 30 (2), 79–88.
- Goel, V., Grossmann, I.E., 2004. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. *Comput. Chem. Eng.* 28 (8), 1409–1429.
- Goel, V., Grossmann, I.E., 2006. A class of stochastic programs with decision dependent uncertainty. *Math. Program.* 108 (2–3), 355–394.
- Gunasekaran, A., Ngai, E.W., 2012. The future of operations management: an outlook and analysis. *Int. J. Prod. Econ.* 135 (2), 687–701.
- Hällgren, M., Maaninen-Olsson, E., 2005. Deviations, ambiguity and uncertainty in a project-intensive organization. *Proj. Manag. J.* 36 (3), 17–26.
- Hansen, M.J., Vaagen, H., van Oorschot, K., 2020. Team collective intelligence in dynamically complex projects? a shipbuilding case. *Proj. Manag. J.* 8756972820928695.
- Hazir, O., Ulusoy, G., 2019. A classification and review of approaches and methods for modeling uncertainty in projects. *Int. J. Prod. Econ.* 107522.
- Herroelen, W., Leus, R., 2005. Project scheduling under uncertainty: survey and research potentials. *Eur. J. Oper. Res.* 165 (2), 289–306.
- Herroelen, W., Leus, R., Demeulemeester, E., 2002. Critical chain project scheduling – do not oversimplify. *Proj. Manag. J.* 33 (4), 46–60. URL: <https://lirias.kuleuven.be/handle/123456789/122992>.
- Huchzermeier, A., Loch, C.H., 2001. Project management under risk: using the real options approach to evaluate flexibility in R...D. *Manag. Sci.* 47 (1), 85–101.
- Ibadov, N., Kulejewski, J., 2019. Construction projects planning using network model with the fuzzy decision node. *Int. J. Environ. Sci. Technol.* 16 (8), 4347–4354.
- Jonsbråten, T.W., Wets, R.J.-B., Woodruff, D.L., 1998. A class of stochastic programs with decision dependent random elements. *Ann. Oper. Res.* 82, 83–106, 0.
- Jørgensen, T., Wallace, S.W., 2000. Improving project cost estimation by taking into account managerial flexibility. *Eur. J. Oper. Res.* 127, 239–251.
- Kadane, J., Wolfson, L.J., 1998. Experiences in elicitation. *J. Roy. Stat. Soc.: Series D (The Statistician)* 47 (1), 3–19.
- Khodakarami, V., Abdi, A., 2014. Project cost risk analysis: a bayesian networks approach for modeling dependencies between cost items. *Int. J. Proj. Manag.* 32 (7), 1233–1245.
- King, A.J., Wallace, S.W., 2012. Modeling with stochastic programming. In: Springer Series in Operations Research and Financial Engineering. Springer.
- King, A.J., Wallace, S.W., Kaut, M., 2012. Scenario-tree generation. In: King and Wallace (2012). Ch. 4.
- Lambrechts, O., Demeulemeester, E., Herroelen, W., 2010. Time slack-based techniques for robust project scheduling subject to resource uncertainty. *Ann. Oper. Res.* 186 (1), 443–464.
- Lium, A.-G., Crainic, T.G., Wallace, S.W., 2007. Correlations in stochastic programming: a case from stochastic service network design. *Asia Pac. J. Oper. Res.* 24 (2), 161–179.
- Lium, A.-G., Crainic, T.G., Wallace, S.W., 2009. A study of demand stochasticity in stochastic network design. *Transport. Sci.* 43 (2), 144–157.
- Loch, C.H., DeMeyer, A., Pich, M., 2011. Managing the Unknown: A New Approach to Managing High Uncertainty and Risk in Projects. John Wiley & Sons.
- Luh, P.B., Liu, F., Moser, B., 1999. Scheduling of design projects with uncertain number of iterations. *Eur. J. Oper. Res.* 113 (3), 575–592.
- Neumann, K., 1990. Stochastic Project Networks: Temporal Analysis, Scheduling and Cost Minimization, vol. 344. Springer Science & Business Media.
- Nichols, K., 1990. Getting engineering changes under control. *J. Eng. Des.* 1 (1), 5–15.
- Özdamar, L., Alanya, E., 2001. Uncertainty modelling in software development projects (with case study). *Ann. Oper. Res.* 102 (1–4), 157–178.
- Petit, Y., Hobbs, B., 2010. Project portfolios in dynamic environments: sources of uncertainty and sensing mechanisms. *Proj. Manag. J.* 41 (4), 46–58.
- Pich, M.T., Loch, C.H., Meyer, A.D., 2002. On uncertainty, ambiguity, and complexity in project management. *Manag. Sci.* 48 (8), 1008–1023.
- Schuyler, J.R., 2001. Risk and Decision Analysis in Projects.
- Servranckx, T., Vanhoucke, M., 2019. A tabu search procedure for the resource-constrained project scheduling problem with alternative subgraphs. *Eur. J. Oper. Res.* 273 (3), 841–860.
- Simchi-Levi, D., 2010. Operations Rules: Delivering Customer Value through Flexible Operations. MIT Press.
- Simchi-Levi, D., Schmidt, W., Wei, Y., Zhang, P.Y., Combs, K., Ge, Y., Gusikhin, O., Sanders, M., Zhang, D., 2015. Identifying risks and mitigating disruptions in the automotive supply chain. *Interfaces* 45 (5), 375–390.
- Vaagen, H., Kaut, M., Wallace, S.W., 2017. The impact of design uncertainty in engineer-to-order project planning. *Eur. J. Oper. Res.* 261 (3), 1098–1109.
- Vaagen, H., Wallace, S.W., 2008. Product variety arising from hedging in the fashion supply chains. *Int. J. Prod. Econ.* 114 (2), 431–455.
- Vaagen, H., Wallace, S.W., Kaut, M., 2011a. Modelling consumer-directed substitution. *Int. J. Prod. Econ.* 134 (2), 388–397 special issue on 'Robust Optimization in Supply Chain Management'.
- Vaagen, H., Wallace, S.W., Kaut, M., 2011b. The value of numerical models in quick response assortment planning. *Prod. Plann. Contr.* 22 (3), 221–236.
- Van de Vonder, S., Demeulemeester, E., Leus, R., Herroelen, W., 2006. Proactive-reactive project scheduling trade-offs and procedures. In: Józefowska, J., Weglarz, J. (Eds.), Perspectives in Modern Project Scheduling. Of International Series in Operations Research & Management Science, vol. 92. Springer, pp. 25–51.
- Vanhoucke, M., Coelho, J., 2016. An approach using sat solvers for the rcpsp with logical constraints. *Eur. J. Oper. Res.* 249 (2), 577–591.
- Wall, D.M., 1997. Distributions and correlations in Monte Carlo simulation. *Construct. Manag. Econ.* 15 (3), 241–258.
- Ward, S., Chapman, C., 2003. Transforming project risk management into project uncertainty management. *Int. J. Proj. Manag.* 21 (2), 97–105.
- Zhu, G., Bard, J.F., Yu, G., 2005. Disruption management for resource-constrained project scheduling. *J. Oper. Res. Soc.* 56 (4), 365–381.