ABSTRACT

Title of dissertation: OPTIMAL REQUIREMENT DETERMINATION
FOR PRICING AVAILABILITY-BASED
SUSTAINMENT CONTRACTS

Amir Reza Kashani-Pour, Doctor of Philosophy, 2017

Dissertation directed by: Professor Peter A. Sandborn
Department of Mechanical Engineering

Sustainment constitutes 70% or more of the total life-cycle cost of many safety-, mission- and infrastructure-critical systems. Prediction and control of the life-cycle cost is an essential part of all sustainment contracts. For many types of systems, availability is the most critical factor in determining the total life-cycle cost of the system. To address this, availability-based contracts have been introduced into the governmental and non-governmental acquisitions space (e.g., energy, defense, transportation, and healthcare). However, the development, implementation, and impact of availability requirements within contracts is not well understood.

This dissertation develops a decision support model based on contract theory, formal modeling and stochastic optimization for availability-based contract design. By adoption and extension of the "availability payment" concept introduced for civil infrastructure Public-Private Partnerships (PPPs) and pricing for Performance-Based Logistics (PBL) contracts, this dissertation develops requirements that maximizes the outcome of contracts for both parties.

Under the civil infrastructure "availability payment" PPP, once the asset is available for use, the private sector begins receiving a periodical payment for the contracted number of years based on meeting performance requirements. This approach has been applied to

highways, bridges, etc. The challenge is to determine the most effective requirements, metrics and payment model that protects the public interest, (i.e., does not overpay the private sector) but also minimizes that risk that the asset will become unsupported. This dissertation focuses on availability as the key required outcome for mission-critical systems and provides a methodology for finding the optimum requirements and optimum payment parameters, and introduces new metrics into availability-based contract structures.

In a product-service oriented environment, formal modeling of contracts (for both the customer and the contractor) will be necessary for pricing, negotiations, and transparency. Conventional methods for simulating a system through its life cycle do not include the effect of the relationship between the contractor and customer. This dissertation integrates engineering models with the incentive structure using a game theoretic simulation, affine controller design and stochastic optimization. The model has been used to explore the optimum availability assessment window (i.e., the length of time over which availability must be assessed) for an availability-based contract.

OPTIMAL REQUIREMENTS DETERMINATION FOR PRICING
AVAILABILITY-BASED SUSTAINMENT CONTRACTS

by

Amir Reza Kashani-Pour

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in Partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:

Professor Peter Sandborn, Chair
Professor Jeffrey Herrmann
Associate Professor Qingbin Cui
Assistant Professor Huan Xu
Assistant Professor Monifa Vaughn-Cooke

Dedication

To America!

and to you as a reader.

Acknowledgments

What a ride! My years at the University of Maryland, College Park have exposed me to a variety of challenging, invigorating and enjoyable experiences, and I would like to take this opportunity to thank all the excellent professors, fantastic staff, and great colleagues whom I have been fortunate to get to know and learn a lot from them.

I would like to sincerely thank and express my highest gratitude to my advisor Prof. Peter A. Sandborn for giving me the opportunity to work with him on such challenging and fascinating research problems. I'd be nowhere near the researcher I am today without him. His vision, insightful comments, careful examination, and mentorship helped me to accomplish many research goals throughout this dissertation. I hope that this dissertation exemplifies his philosophy of conducting high-impact research that addresses complex questions head-on and results in solutions to public sector challenges around the globe. My gratitude to Professor Sandborn for his endless patience, energy, and unstoppable support during my research endeavor will remain with me in all future endeavors.

I am grateful to Prof. Quinbin Cui for introducing me to public-private partnerships, which added a new dimension to the methodology of this work. And many thanks to Prof. Jeffrey W. Herrmann, Dr. Xu, and Dr. Vaughn-Cooke for serving as my committee members, and contributing their time, feedback, and helpful feedback. I am grateful to Mr. William Lucyshyn for his feedback and his support as an extended committee member.

I would like to recognize the many professors at UMCP who taught me so much about reliability modeling, simulation, optimization, statistical analysis, and federal acquisitions, including Prof. Michael Fu, Benjamin Kedem, Mark Austin, Shapour Azarm, Mohammad Modarres and the Honorable Jacques Gansler. I need to also thank Professor Stephen P. Boyd from Stanford University, for generously providing CVX free, open-source and along with great teaching material. For many great discussions, collaborations and the friendships, I am thankful to my fellow graduate student researchers in the Center for Advanced Life Cycle Engineering (CALCE) including, but not limited to, Roozbeh Bakhshi, Hannes Greve, Xin Lei, Navid Goudarzi, Helmut Bevensee, Nathan Valentine, Mehdi Kohani, Chandradip Patel, Thomas Fritzler, Anshul Shrivastava, Arvind SSV, Rayhoo Zhang, Taoufik Jazouli, Gilbert Haddad and Jordan Jameson. It has been a pleasure to work with you all towards our collective goal of understanding and improving the reliability of many different systems to support sustainable and dependable technological growth.

I could not have pursued my dreams to this point, without the supports, sacrifices, and love from my parents (Ahmad and Nasrin), truly inspiring brothers(Alireza and Mahdi), and my amazing Christine. I am thankful to many of my inspiring and supportive teachers

who I keep what I learned from them close to my heart. Finally, I am grateful to the only grandfather I had, as he passed away while I was buried under my research work and VISA limitations and could not say a proper farewell to him. I am thankful to him for all the peace he brought to my family, his curiosity, persistence, and passion for lifelong learning and re-examining. I also have to recognize the role of my great uncle, Ebi, his family and their love of mind-blowing books, for helping me discover much more than I could ever afford in any part of my formal education. I am thankful to my other supportive uncle, Mahmood, for generously helping me with my big trip to the US.

I am indebted to my *barobach* from the Iranian Graduate Student Foundation (IGSF), colleagues in the Mechanical Engineering Graduate Association (MEGA) and fellow representatives in Graduate Student Government (GSG) for all their volunteering efforts in improving the sense of support and community on the campus.

During my years in DC metro area, I shared many great moments with exceptional people, whom I would like to take the chance to acknowledge for all great memories which are now a major part of my "graduate student life." I wish there were enough space to thank you all one by one, but I know that you know who you are, especially if you are reading this!

Contents

List of Tables

List of Figures

List of Abbreviations

A_m, A_o	Materiel availability and operational availability
b	Back-order rate per for unit of repairable item
Budget(t)	Budget at time t
C	Cost
C_a	Annualized cost
C_m	Total inventory cost
C_o	Operating cost
C_c	Cost of operation to customer
C_s	Cost of operation to contractor
DT	System downtime
d	Effective discount/interest rate
$F_x(t)$, $f_x(t)$	Cumulative distribution and probability density functions
i	Index for counting
$c(p)$	Operation downtime penalty per unit of time
$C_3(t_p)$, $C_4(t_r)$	Costs for repairing a part and costs for replacing a part
h	Holding cost per unit
k	Index for counting
$P(z)$	Probability of z
MAP(t)	Maximum availability payment
$m(u)$	Renewal density function of the system
MTFBC	Mean-time-between-failures in calendar time
n	Index

$N(\mu, \sigma^2)$	Normal distribution with mean μ and variance σ^2
N_a	Number of assessments during the contract time
$N_f(\Delta t)$	Number of failures in timer period of $\Delta t = t_2 - t_1, t_2 > t_1$
N_s	Number of systems in a fleet
N_d	Number of days in each assessment
N_a	Number of performance assessments per contract
$R(t)$	Reliability function
s	Shipping spare part unit cost
t	Time index
T_t	Specific time
TTR	Time for performing a repair-by-replacement job
t_m	Maintenance time
t_p	Cycle time for replacing a part
t_r	Cycle time for repairing a part
T	Contract length
T_a	Assessments interval in unit of time
UP	System operating time (uptime)
$u(t)$	Order size at time t
$w(t)$	Demand for inventory at time t
$x(t)$	Inventory level at time t
$y(t)$	Performance measure
$v(t)$	Availability performance measure
$W(\alpha, \beta)$	Weibull distribution with shape α, scale β
α	Inherent Weibull shape
α_p	α-service level (type I service level)
β	Inherent Weibull scale parameters
β_p	β-service level (type II service level)

η	Bankruptcy inhibition constraint
λ_p	Planned replacement rate under
μ	Mean
ϱ	Correlaton
σ	Standard deviation

Chapter 1:   Introduction

Understanding the total life-cycle cost is an essential part of pricing for any procurement/sustainment acquisition or service contract. For many safety-, mission- and infrastructure-critical systems, availability is an important factor in determining the life-cycle cost. Common wisdom is that higher reliability and more efficient supply-chain management improves the availability of systems; however, it is also important to explore how availability drives system and supply-chain attributes. For example, how can the contractor establish efficient and cost-effective management approaches given specific availability requirements? What are the methodologies and quantitative methods for designing an availability requirement in an optimum way? This chapter addresses the current state of knowledge by reviewing: 1) the definitions and existing approaches used to design and plan for availability contracts; and 2) the associated approaches for decision modeling for sustainment acquisition. Chapter 2 addresses the challenges, gaps and opportunities for new research in this area, followed by the formal problem statement.

1.1   Introduction to Availability-Based Contracts

A significant shift toward a service-based economy has forced organizations to modify their business philosophy from product-centric to service-oriented through out-sourcing logistics and maintenance (Baines et al., 2009). Outcome-based contracts that pay for effectiveness and penalize performance shortcomings have been introduced and referred to as pay-per-hour or performance-based contracts, e.g., performance-based logistics (PBL) contracts used by the United States Department of Defense (DoD), and "availability contracting

maintenance model" initiated by Rolls-Royce (Bangemann et al., 2006). These concepts are also being used for federal acquisition of healthcare (Eijkenaar et al., 2013), energy, infrastructure, and in other sectors. Outcome-based contracts allow customers to pay only for the specific outcomes achieved rather than workmanship and materials delivered. One of the merits of outcome-based contracts is the optimally sharing of risks by both parties. These contracts present a pricing challenge due to a dramatic alteration of the risk sharing scheme when compared to conventional contracts like fixed-price or cost-plus (Kim et al., 2007). Underestimating the risks involved and therefore the contract cost have caused some projects to stop and given rise to doubts about the applicability of this class of policy for new acquisition contracts (Thompson, 2010). The challenge also exists for designing availability requirements given uncertainties in both inputs and outcomes over a long time of period. This issue has been largely ignored for a several reasons: 1) availability-based contracts are a relatively new concept and there is not enough historical data to evaluate their effectiveness and the success of contractors that implement them; 2) the engineering design process does not directly target the availability (or other contractual outcomes), but rather focuses on immediate preferences like performance, purchase price; 3) the logistics and maintenance of the system are planned and executed separately from the design of the system (often as an afterthought); 4) design methodologies are not equipped to handle availability and other outcomes as design inputs; and 5) usage behavior and incentives for maintenance contractors are being neglected in most design cases due to the high level of uncertainty in the operational phase of the life cycle. An integrated approach that includes supply chain, inventory and maintenance management to enable direct evaluation of different contracting and support policies is needed.

The next section provides a simple qualitative example that demonstrates how an availability-based contract could work (many variations are possible and in use today). It describes who the parties to the contract are and how the contract agreement is transacted between them. After the example, the remainder of this chapter describes and defines the

main elements of availability-based contracts from the viewpoint of the contractor and the customer. It explains the most important factors involved in designing availability-based contracts and how this class of contracts can effectively improve the preparedness and cost effectiveness of sustainment activities if the design issues are addressed.

1.2 *Availability Contract Examples*

This section focuses on a simplistic case that addresses the practical aspects of availability-based contracts. We assume that there is a costumer that uses parts from an inventory and the contract governs the availability level of the inventory of these parts, where the inventory is operated (managed) by a contractor (who is not the customer).

In the simplest version of this problem, we assume that the objective of the contractor is to minimize their costs. We also assume that the customer owns the inventory facility and the contractor will pay for the use of the space that the parts inventory occupies.[1] The availability contract articulates a penalty that is imposed on the contractor if the availability of the inventory drops below a certain level. As described, this example is a simplified version of Performance Based Logistics (PBL) that is currently implemented by the US DoD. In our simple example, we do not consider any profit-sharing between the contractor and the customer, but interested readers should see (Hamidi et al., 2014).

In order for the customer to enforce the inventory availability requirement in the contract, there have to be assessments. An assessment is defined as checking the performance of the contractor based on the predefined metrics (e.g., reliability, inventory level and back-orders) and we assume that at each assessment point the customer can terminate the contract if the performance of the contractor does not meet the criteria specified in the contract.

---

[1] In this simple example we will assume that the contractor buys the parts (i.e., the inventory) from an OEM and the contractor owns the parts forever (even after the customer takes the parts from the inventory and puts them into their system). For real-world contracts, the actual ownership of the inventory depends on the particular situation and the contract.

We assume that during the period of contract (the length of time covered by the contract), there will be N_a assessments. At each assessment the performance of the contractor will be measured; based on the results of the assessment, the customer's payment to the contractor could be adjusted (penalized or incentivized).[2] Therefore, the contractor needs to adjust their planning horizon around the periods associated with the assessments.

Everything in this contract costs the contractor money. The contractor must pay to hold the parts in inventory, order additional parts, tie up capital in the parts, for delays caused by back-orders, and the penalties assessed on the contractor by the customer. But it will be the customer that pays for the cost of assessing the contractor performance (GAO-02-1049, 2002; GAO-05-966, 2005). This assessment can be done on a regular basis or randomly by a performance review unit consisting of both parties (e.g., performance review board).

This example is a simplified version of an availability-based contract, many PBL contracts operate similarly to this example. There are many variations of this in use today and just about every availability-based contract is unique (Gansler & Lucyshyn, 2006). In the next subsection, a particular availability contract is described.

1.2.1 Michelin Tire/US Navy Contract

In this subsection, we describe a a real-world example of availability-based contracting. The example is a contract between US Navy (the customer) and Michelin Aircraft Tire Co, LLC (MATC) (the contractor) associated with sustaining the aircraft tires used by U.S. Navy over a wide range of aircraft including F-14, F-18, and A-4. MATC uses Lockheed Martin as a sub-contractor for supply-chain management. The goal of the US Navy is to improve the availability (more exactly, the fill-rate) while reducing the inventory level. The availability-based contract is used to guarantee aircraft fleet readiness and reduce the logistics cost to the customer. The primary metrics are the fill-rate and reliability of the tires. The decision to repair/replace is done based on (MIL-PRF-7726K, 2007) for repairable tires

---

[2]In some PBL contracts, the customer may also have the flexibility to cancel the contract, replace the contactor with another contractor, or take over the operation themselves.

and (MIL-PRF-5041K, 1998) for new tires. The contract requires that MATC fills 95% of the tire requisitions within 2 working days in the continental United States and 4 working days outside of the continental United States. The length of the contract is five years with two five-year extension options. Fill-rate is defined as the number of requisitions filled within the time criteria divided by the total number of requisitions during a measurement period. Measurement periods are defined as increments of 6 months starting at the contract award.

In terms of the scope of operation of the contractor, the "baseline requirements" of the Michelin contract asks for an "Intermediate to Depot" maintenance concept. Maintenance operations of an organization are generally divided into three types: 1) organizational-level maintenance, which are maintenance activities performed directly on a system or its support equipment (e.g., repair, inspection, testing or calibration); 2) intermediate-level maintenance, which is done on removed component parts or equipment at a "shop"; and 3) depot-level maintenance, which is done at a major repair facility (Dulcos & Shepherd, 1991). If a malfunction is diagnosed in a system, the malfunctioning item is removed from the system and brought to the base supply. If a spare is available it is installed in the system; otherwise a back-order is established for that item. Because this item is directly installed in the system, the back-order implies that there is a "hole" in the system that causes it to be non-operational. In our example the fill-rate is measured at the intermediate to depot level and the back-order does not imply a "hole" in the system, but rather a shortage of parts in the inventory.

The payments in this contract are based on Annual Firm Fixed Pricing for Level of Performance addressing operations to support a forecasted demand over a fifteen-year period to include a five-year base period, and two five-year options. The contractor bills the Government monthly for $1/12^{th}$ of the estimated annual cost.

One significant challenge in these availability-contracts is that the period of perfor-mance (the assessment period) is arbitrary, i.e., not based on any carefully constructed

analysis, but rather probably determined by convenience or based on what was done in previous contracts. For example in Section 1.2.1 the assessment period is bi-annual. This dissertation addresses the optimization of the period of performance assessment in these contracts.

1.3   Desired Outcome/Performance Measure: Availability

Outcomes can be divided into performance and availability. Figure 1. The concept of availability is important as it accounts for both the frequency of the failure (reliability) and the ability to restore the service or system to operation after a failure (maintainability). In industries with complex systems for which product downtime has a very high cost, availability is often the single most driver of total life-cycle cost. The maintenance ramifications translate into how quickly the failure can be isolated, and the system can be repaired and/or restored. These tasks are usually driven by fault detection, isolation and prognosis followed by inventory response.
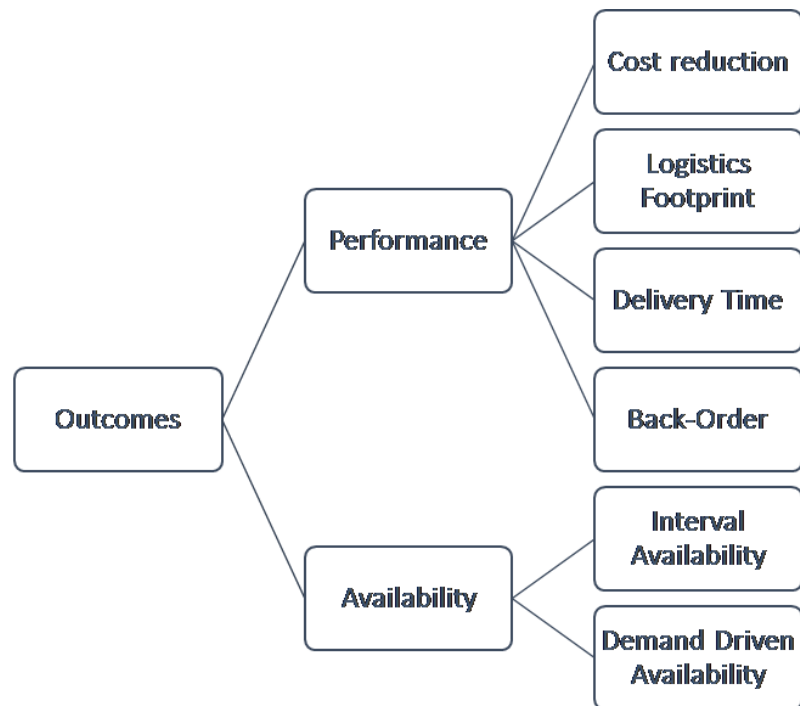
Figure 1: Categorization of contractual outcomes

Availability is the probability that the system is operating properly when it is required for use. In other words, availability is the probability that a system is not failed or undergoing a repair action when it needs to be used. The definition of availability is somewhat flexible, depending on what types of downtimes are considered in the analysis. As a result, there are a number of different types of availability, point (instantaneous) availability, mean availability, steady-state availability, operational availability, interval availability, and materiel (inventory) availability, network availability, fleet availability and layered availability. It needs to be noted that availability-based contracts do not differentiate between sources of success in the outcomes making it difficult to identify the effect of using such contracts on reliability improvement (Newsome, 2008; U.S. Government Accountability Office, 2008).

1.3.1 Point Availability

Point (or instantaneous) availability is the probability that a system (or component) will be operational at any time, t. Point availability is similar to reliability in that it gives the probability that a system has no failures in the interval from 0 to t. Unlike reliability, the point availability measure incorporates maintainability information. At any given time t, the system will be operational if the following conditions are met: it functioned properly during the time interval from 0 to t with probability $R(t)$, and, it functioned properly since the last repair at time u, $0 < u < t$, with probability:

$$\int_0^t R(t-u)m(u)\,du \tag{1}$$

in which $R(t)$ is the reliability of the system at time t and $m(u)$ is the renewal density function of the system. The point availability is the summation of these two probabilities, or:

$$A(t) = R(t) + \int_0^t R(t-u)m(u)\,du \tag{2}$$

$$A(t) = R(t) + R(t - u) \tag{3}$$

1.3.2 Mean Availability

The mean availability is the proportion of time during a mission or time-period that the system is available for use. It represents the mean value of the point availability function over the period $(0, T_t)$:

$$A_e(T_t) = \frac{1}{T_t} \int_0^{T+t} A(t) \, dt \tag{4}$$

1.3.3 Steady-State Availability

The steady-state availability of the system is the limit of the point availability function as time approaches infinity:

$$A(\infty) = \lim_{T_t \to \infty} A(T_t) \tag{5}$$

1.3.4 Operational Availability

Operational availability is a measure of the availability that includes all experienced sources of downtime, such as administrative downtime, logistic downtime, etc. The equation for operational availability is:

$$A_o = \frac{\text{Uptime}}{\text{Operating Cycle}} \tag{6}$$

where the operating cycle is the overall period, time, term of operation being investigated and uptime is the total time the system was functioning during the operating cycle. When there is no logistic downtime or preventive maintenance specified, the operational availability equation returns the mean availability of the system. The system's availability measure approaches the operational availability as more sources of downtime are specified, such as crew logistic downtime, spares logistic downtime, restock logistic downtime, etc. In all other cases, the availability measure is the mean availability. Note that the operational availability is the availability that the customer experiences. It is essentially a posteriori

availability based on actual events that happened to the system. Operational availability can be effectively estimated by accumulating times in discrete-event simulators. The previous availability definitions are priori estimations based on models of the system failure and downtime distributions. In many cases, operational availability cannot be controlled by the manufacturer due to variation in location, resources and other factors that are the sole province of the end user of the product.

1.3.5   Interval Availability

Interval availability is the fraction of time the system is operational during a given interval of time. When there the focus is on the transient behavior of a system or continuous demand, the interval availability is a relevant measure. For instance, the amount of crude oil or natural gas to be delivered over a finite period requires related platforms to be available for a certain number of hours in that window. Figure 2 shows that depending on the interval; the availability can vary greatly within the intervals, while still resulting in the same overall availability.
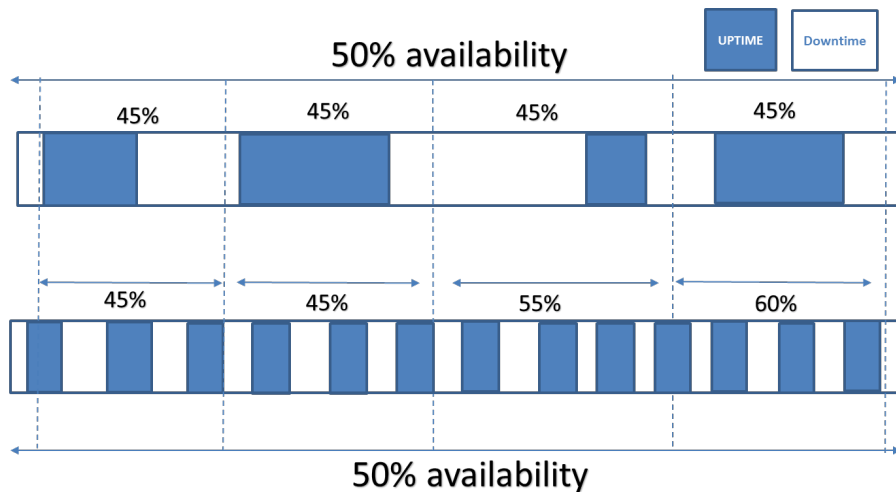
Figure 2: Interval availability with different intervals

1.3.6   Materiel Availability

Materiel availability is a measure of the fraction of the total inventory of a system operationally capable of performing (ready for tasking) an assigned mission at a given time, based on the materiel condition. Materiel availability can be expressed mathematically as (the number of operational end items divided by the total population).[3] Determining the optimum value for materiel availability requires a comprehensive analysis of the system and its planned use, including the planned operating environment, operating tempo, reliability alternatives, maintenance approaches, and supply-chain solutions. Materiel availability is primarily determined by system downtime, both planned and unplanned, requiring the early examination and determination of critical factors, such as the total number of end items to be fielded and the major categories and drivers of system downtime. The materiel availability key performance parameter must address the total population of end items planned for operational use, including those temporarily in a non-operational status once placed into service (such as for depot-level maintenance).

Materiel availability can be expressed in different ways. The following definition represents a point (instantaneous) estimate for materiel availability as a measure, expressed as a fraction of systems (end items).

$$\text{Materiel Availability } (A_m) = \frac{\text{Number of Operational End Items}}{\text{Total Population}} \tag{7}$$

The key elements that must be incorporated in any assessment of A_m are: any measure of A_m must include the total population of systems (end items) to be fielded; any measure of A_m must consider the total life-cycle timeframe of the system (end item); and any measure of A_m must include all major categories of downtime, both planned and unplanned. These are the distinguishing features of the materiel availability metric that differentiate it from the more familiar operational availability metric (uptime/uptime downtime).

---

[3]Note, this is the same definition as "yield", however, yield refers to the outcome of a manufacturing process.
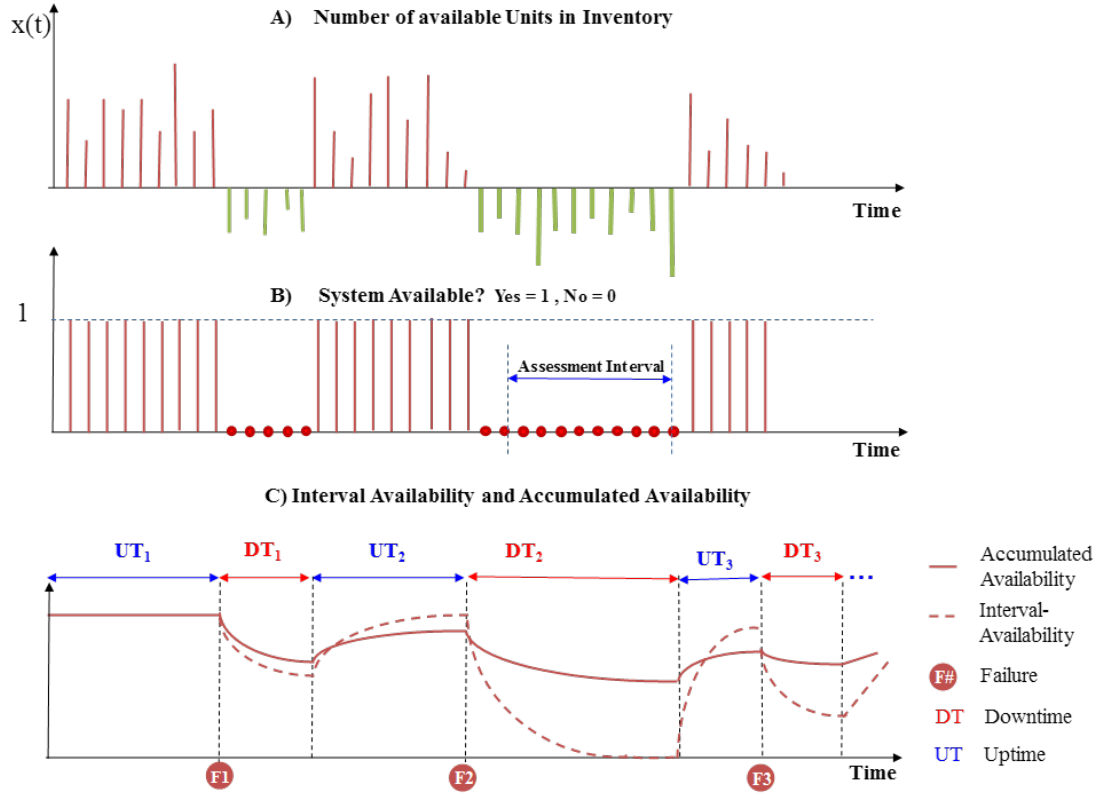
Figure 3: Materiel availability, state of the system & interval availability

As the Figure 3 depicts, it is clear how A_m differs from A_o as it applies to the number of units in the entire fielded inventory of systems, over the entire life cycle of the system and incorporates all categories of downtime. In fact, uptime and downtime of an inventory can be defined based on the level of available inventory or materiel availability as well. However, the best way to view the relationship between A_m and A_o is to see A_m as a function of A_o, together with many other variables. The best way to assess both A_m and A_o is through comprehensive modeling and simulation. Materiel reliability is the cornerstone that insures both A_m and A_o requirements can be met. A_m is far more important in determining the level of availability that is achievable than any other component of logistics system ("Materiel Availability", 2010).

1.3.7 Service level requirements

Service level is commonly used in supply chain management and inventory management to measure the performance of inventory replenishment policies. Several definitions of service levels are used in the literature as well as in practice. We introduce the two main definitions that are related *to Materiel Availability*.

α-service level (type I)

The α-service level is an event-oriented performance criterion. It measures the probability that all customer orders arriving within a given time interval will be completely delivered from stock on hand, i.e. without delay.

Two versions are discussed in the literature differing with respect to the time interval within which the orders arrive. With reference to a demand period, α denotes the probability that an arbitrarily arriving customer order will be completely served from stock on hand, i.e. without an inventory-related waiting time:

$$S_\alpha = \Pr(\text{Period Demand} < \text{Inventory on hand at the beginning of period})$$

In order to determine the safety stock that guarantees a target service level, the stationary probability distribution of the inventory on hand must be known. This version of α is also called the ready rate.

β−service level (type II)

The β-service level is a quantity-oriented performance measure describing the proportion of total demand within a reference period that is delivered without delay from stock on hand:

$$S_\beta = 1 - \frac{\text{Expected backorder per time period}}{\text{Expected period Demand}}$$

This is equal to the probability that an arbitrary demand unit is delivered without delay.

Because, contrary to the variations of the α-service level, the β-service level does not only reflect the stockout event but also the amount backordered, it is widely used in industrial practice. For example, if customer orders total 1000 units, and you can only meet 900 units of that order, your β-service level is 90%. It is also being called the fill rate.

The time-period in this definition is assumed to be sufficiently large enough to capture a single order cycle. For example, a monthly fill-rate can be defined as:

$$\text{Monthly Fill-rate} = \frac{\text{Quantity Ordered per Month} - \text{Quantity Backordered per Month}}{\text{Quantity Ordered per Month}}$$

Also, by the definition, $S_\alpha \leq S_\beta$ whenever the probability of zero demand equals 0.

1.3.8 Network Availability, Fleet Availability, Layered Availability

Network services are distributed across several nodes and can depend on the performance of each node and the demand on each node. Defining and measuring availability requirements for networked systems or fleets in a way that satisfies the ultimate goal of the customer (e.g., preparedness) is not trivial (Immonen & Niemelä, 2008). Difficulties defining and measuring availability metrics makes it even more difficult to predict the availability over a fleet or network of subsystems and parts (Mickens & Noble, 2006). Targeting availability in different layers of the system increases the dimension of the problem, and this problem commonly is called fleet availability.

In summary, one should ask why availability is being used in performance-based contracts instead of other performance metrics or why it is being distinguished as a special performance metric from other metrics? Availability along with technical performance, cost and process efficiency are key elements of support effectiveness. Availability is often an observable and measurable index, especially for a combination of multiple complex subsystems (i.e., system of systems). One can combine the availability of different subsystems

to achieve the availability of the platform and the reverse can also be done under certain conditions. In service contracts for complex systems, availability may be a less ambiguous factor to rely upon to reward the effectiveness of the efforts of the contractors. Availability is also directly connected to reliability and the quality of support, whereas performance is highly affected by the users of the system and the engineering design philosophy. Lastly, when aiming to maintain fleet-level preparedness, availability is more closely connected to sustainment contractors.

1.4   Performance Evaluation and Performance Sampling Procedures

In availability-based maintenance contracts, customers (e.g., road administrators, the DoD, etc.) define performance measures that specify the minimum condition to which the asset items must be maintained. To ensure that contractors maintain the asset items according to these measures, customers must design and implement a comprehensive and reliable performance monitoring process.

One of the most important areas within the performance monitoring process is inspection conducted in the field. Defining a procedure that guarantees the success of field inspections is a challenge. There are generally two categories of performance sampling: deterministic and statistical.

Deterministic sampling involved looking at the performance at the end of some contractual period and determining the payment based on the relationship between payment and performance. In statistical sampling, for example in a performance-based maintenance of a highway, the customer can only sample the quality of pavement a few times and in a few places (Ozbek & Jesus, 2007). Based on this concept, Pinero (2003) developed a statistical sampling procedure to ensure that findings from field inspections will be reliable and representative, with high confidence, of the condition of asset items in the entire population.

1.4.1 *Analysis and Modeling of Time-Correlated Failures*

The majority of existing reliability-based work assume failures are identically, and independently distributed. This assumption does not take into account the time-varying behavior of failures, the periodic behavior of failures and peak periods in the number of failures over time(Carroll et al., 2015). The presence of time correlations between failures including periods with increased failure rate, rejects this assumption and can have a significant impact on the effectiveness of the maintenance optimization, or fault-tolerance strategy (Yigitbasi et al., 2010). Understanding the temporal correlations and exploiting them for optimum checkpointing and scheduling decisions provides new opportunities for enhancing conventional maintenance optimization and contract design.

1.5 *Contract Mechanism: Performance-Based Contracting (PBC)*

Performance-based contracting (also referred to as performance-based life cycle product support and performance-based logistics (PBL)) refers to a group of strategies for system support that instead of contracting for goods and services, the contractor delivers performance outcomes as defined by performance metric(s) for a system or product. PBC thinking is reflected in a famous quote from Theodore Levitt (Levitt, 1972): "The customer does not want a drilling machine; he wants a hole-in-the-wall." PBC and similar outcome-based contracts (Table 1.1) pay for effectiveness (availability, readiness and/or other related performance measures) at a fixed rate, penalize performance shortcomings, and/or award gains beyond target goals. Table 1.1 describes outcome-based contracts in terms of incentives and payment models.

PBL is the purchase of support as an integrated, affordable, performance package designed to optimize system readiness and meet performance goals for the system through long-term support.

The top-level metric objectives are operational availability, reliability, cost per unit,

Table 1.1: Contract Types Based on Incentive Structure

Contract Type	Criteria	Taxonomy
Conventional Contract (Labor/Material/Hour)	This is a conventional contract, the customer pays for the cost of material, labor and hours of performance that the contractor reports	There is no incentive for contractor to optimize their operation or minimize their costs.
Firm Fixed Price	Customer pays a set price, regardless of the contractor's costs or efforts. After the contractor achieves the first best cost reduction effort, they are incentivized to reduce cost as much as possible.	The rigidity of this mechanism fails to control costs, or help the contractor to deal with the risks and uncertainties involved, specially in R&D projects, and instead produces expensive legal battles.
Cost-Plus Award Fee (CPAF)	Customer pays a fee including an award amount to motivate contractor to achieve a certain objective. It should be possible to determine the feasible level of objective before the contract.	Supplier tries to increase the cost as much as possible. Predetermined award fee based on judgmental evaluation of the contract.
Cost-Plus Incentive Fee	The legal agreement specifies a target cost, base contractor pay, a formula to be used to figure the incentive bonus, and minimum and maximum limits on the contractor's pay.	The supplier exerts no cost reduction, and is indifferent to the award. For incentivizing subjective areas of contractor's performance.

logistics footprint and logistics response time. This level of support differs from the 'best effort' approach typical of DoD organic support in terms of having a clear delineation of performance outcome. Under PBL (also called Contract for Availability-CfA), the contractor (system supporter) often commits to providing the current performance level

at a lower cost, or increased performance at a cost similar to that previously achieved under a non-PBL approach (Gruneberg et al., 2007). This concept is known in practice as "Performance Contracting" (Hansen, 2006), "Availability Contracting" (Cushway, 2006), Contract for Availability (CfA) (Hockley et al., 2011), "Performance-Based Logistics" in the defense context and " Performance-Based Service Acquisition (PBSA)" (Gansler, 2000). In the U.S. performance-based logistics is normally established on a contractual basis, whereas in Europe PBL is being categorized under public private partnerships.

Performance-based logistics (PBL) and similar mechanisms have become popular for contracting the sustainment of military systems in the United States and Europe and aim to replace traditional fixed-price and cost-plus contracts to improve product preparedness and reduce the total cost of ownership of systems. PBL has become the US DoD's preferred support strategy for weapons systems (Gansler, 2000). These contracts specify the government's desired result without stipulating how a task should be performed, thus granting contractors the flexibility to complete its tasks in the manner the firm deems most appropriate PBL contracts are normally executed at three levels: component-level, subsystem-level, and system or platform-level. Subsystem-level contracts are the most prevalent form of PBL. In a subsystem PBL contract, the contractor is tasked with sustaining a subsystem over a period of 5-10 years[4] – often the subsystem has previously been supported via a non-PBL contract. Many of today's PBL contracts use what is referred to as public-private partnerships (PPPs). In a subsystem PBL, a PPP could mean that the contractor partners with government owned and staffed maintenance facility. The contractor brings in their best practices and manages the facility, and the contractor is responsible for the outcome.[5]

Several studies have investigated the effectiveness of PBLs on product reliability

---

[4] United States government PBL contracts are limited to a maximum 5 years by law. (FAR 16.505, FAR 17.104, FAR 17.204, FAR 22.1002)

[5] PPPs in the civil infrastructure area (e.g., highway construction and support) have a different structure than those referred to in subsystem PBL. Civil infrastructure PPPs require the private sector to take responsibility for designing, building, financing, operating and maintaining an asset, which is a much broader view than today's subsystem PBL PPPs in use in the U.S. Department of Defense.

17

in aerospace and electronics industry (Hockley et al., 2011; Kim et al., 2007) as well as quantitative survey analysis over different defense agencies (Randall et al., 2011). They have attempted quantitatively to relate the effect of incentives under this contract with reliability improvement under PBL against conventional contracts. The expectation is that the incentives of contractors will drive their course of action in the design process so we should be able to see either an increase in scheduled maintenance or increase of reliability, either way an increase in availability. Under availability-based contracting, manufacturers supposedly move their designs toward higher maintainability and reliability products, which leads to higher availability for the customer. Pricing such activities and progressive decision making throughout the life cycle is a stochastic optimization problem that is being performed today using highly qualitative "fudge factors" (Thompson, 2010). Some authors use game theory to model the interaction between customer and contractor toward decisions to improve the reliability (Ashgarizadeh & Murthy, 2000). Sometimes the customer is interested in layered availability, which means differentiating between subsystems or the interconnection between systems and their interactions. In other cases, the limitation on contractors comes from the fundamental physics and/or properties of the materials and cannot be resolved. Some have argued that these contracts are not necessarily designed to save money, but rather to maintain or improve the current system or platform performance in a cost constrained world (U.S. Government Accountability Office, 2008).

Another significant challenge with PBL contracts is to determine the contract requirements and price that protects the interest of the customer, i.e., which does not overpay the contractor, but also minimizes that risk that the system will become unsupported. Subsystem PBL contracts are generally priced based on: 1) estimating how many units will need repair, 2) how much it will cost for each repair, and 3) how the number of units requiring repair and/or the repair cost will decrease over time as a result of design and/or maintenance improvements made by the contractor. If greater than projected improvements are realized, the money saved is shared with the contractor according to a schedule

negotiated in the contract ("gain share"). Meeting or exceeding target performance may also allow the contractor to add additional years to the contract ("award term"). With subsystem PBL contacts, it is reasonably straightforward for the customer (which is most often the government) to demonstrate a benefit by determining what it would cost to support the system doing business as usual (no improvements, non-PBL contract) compared to the cost of a PBL contract, e.g., often pre-PBL support and performance experience exists. However, for new system acquisition, where there is no sustainment history; and for platform-level PBL, the PBL contract pricing problem is much more complex and it is unclear how to optimally apply PBL contract mechanisms. For example, a recent study of PBL effectiveness (Boyce & Banghart, 2012), reported on the cost of 21 PBL contracts where in 9 out of 9 component and subsystem-level PBL contracts the cost decreased, but for platform-level (called system-level in the study), PBL 6 out of 12 contracts resulted in either cost increases or indeterminable cost changes.

PBL and availability payment PPPs share many characteristics. In both cases, the public and private sector objectives are aligned towards ensuring better value for the end users/public. These contracts are long-term in nature and demand the private sector to play a major role in meeting the objectives of the system or project. The private sector bears the majority of project or system risks and is encouraged to pursue innovative processes and methods. Table 1.2 summarizes the similarities and differences between these contracts. Although the procurement contracts are operated by different public agencies and targeted on different assets, they all must be well-designed and priced to ensure adequate protection of the public interest. In the defense industry, the challenge becomes much greater considering the complexity and uncertainty of defense acquisition programs. While current practices may be effective at the component and subsystem levels, pricing a PBL contract becomes more difficult for a new system acquisition where no prior estimates of any kind are available. Therefore, developing and introducing innovative methods and best practices in civil infrastructure PPPs have great potential to improve

DoD PBL contract acquisition significantly.

Table 1.2: Mapping of Availability Payment Contracts to PBL Contracts

	DoD PBL Contracts	Availability Payment PPP Contracts
Performance	Availability, reliability, downtime, outcome, variances from goals	Physical and qualitative availability, serviceability, resilience, and others
Incentive	Contractor rewarded for performance exceeding expectations	Typically not used. In some cases, incentives are used for qualify materials up to 5% of total construction cost.
Penalty	Penalized for not meeting performance criteria and non-availability	Penalized for not meeting performance criteria and non-availability
Pricing	Bidding	Engineer estimate and bidding
Value for Money	Benchmarking—compare to non-PBL contracts; market research	Value for money analysis to consider unique characteristics of infrastructure project
Contract Term	Medium to long-term (5 year base contract followed by a 5 year extension)-duration based on regulations.	Long-term (minimum 10 year and maximum 99 years), duration based on the value for money analysis
Renegotiation	Allowed and possible	May be allowed

1.6 The Design Process for Availability-Based Contracts

In the transition from conventional purchase models to a performance-based service contract model, we need to re-examine the design process. Engineering system design and pricing have a bi-directional relationship; one can reach the price from the design parameters or derive the design from a given budget. In the majority of engineering projects, especially the ones dealing with designing new products, designers pursue the former method; whereas in practice, and specifically in designing support and sustainment systems, the flow down of requirements will be determined by considering contractual terms

and budgetary constraints. Breaking down the high-level requirements to lower-level requirements requires considering constraints from lower levels (Kohani & Pecht, 2015). This bi-directional relationship creates new constraints for the contractual oriented design (Sun et al., 2009).

The classical design for procurement and support contracts are a trade-off between the costs of providing high reliability (such that the system lasts longer than support contract lengths or the warranty term) and the opportunity costs of the manufacturer or maintenance parties (Frangopol & Maute, 2003).

In terms of life-cycle cost estimation, the design process is conventionally a point-to-point mapping from the space of the design parameters to the space of structural responses (e.g., total life-cycle cost, availability). In this mapping, each point in the space of the design parameters defines a feasible or non-feasible design structure, and all feasible designs guarantee that the predetermined (contractually obligated) outcome requirement is met (Bakhshi & Sandborn, 2017). Most approaches of this kind require many iterations in the design without any guarantee that the requirement is met (Jazouli & Sandborn, 2010). Also, the problem of uncertainty and unavailable data is adding to this mapping challenge. Möller et al. (2011) tackles this problem by using fuzzy processes to capture uncertainty without depending on statistical data. To assess the robustness and agility of such designs more computationally burdensome analysis is needed. Also, when numerous stochastic factors are present the proposed strategy might not be flexible enough for operational support. The uncertainty in achieving the final design specification might not be acceptable in many cases. This will become even more challenging in the context of life-cycle engineering and the support of critical systems with complex supply chains. Risk-based design literature has addressed a similar problem by considering uncertainty propagation through the process and risk allocation and management. In conclusion, the design for availability based contracts has new dimensions that need to be treated separately and specifically from conventional design.

Chapter 2: Problem Construction and Objectives

The general problem of designing an availability-based contract that can provide a win-win situation by minimizing the cost for the contractor and guaranteeing availability for customer is very general and beyond the scope of this dissertation. This dissertation is limited in terms of scope (the life-cycle stage of the systems under the availability-based contracts) and the parameters that can be designed in the contract and cost structure it targets. This dissertation is also looking a single contract between the contractor and customer and ignores the effect of long series of contracts that can occur under relational contracting (Erkoyuncu, 2011).

Based on Operation of the Defense Acquisition System Instruction (DoD Instruction 5000.02), there are multiple stages of the life cycle that contracts can be used in: 1) material solution analysis, 2) technology development, 3) engineering and manufacturing development, 4) production and deployment and 5) operations and support. In this dissertation, we are focusing on operation and support stage of the life cycle with a focus on sustainment and maintenance activities.

In this dissertation, the contract is defined as a set of requirements with a specific payment structure for a certain level performance for a specific length of time between a customer and a contractor (two parties). The availability-contract design problem is defined as finding the optimum requirement parameters under which the minimum required availability level can be achieved for the customer while respecting cost constraints. In practice, the contract can be designed by the customer or by a third-party who works for the customer under a separate contract.
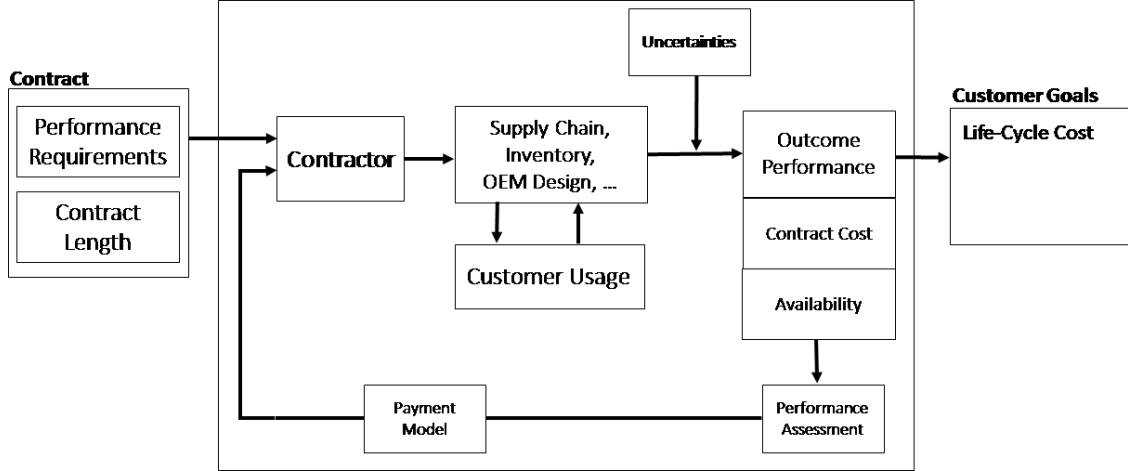
Figure 4: Mapping of the availability-based contract as a control loop

We define the contractual relationship within a control theory framework in which the system and the contractor use a set point defined by customer in a closed-loop control system. Control theory framework can address the dynamics of contractor and performance in time, as well as the impact of monitoring performance on the payments. Figure 4 shows that the contractor's action will be an input to the supply-chain system including the inventory of the sub-assemblies in which their availability is the main factor being monitored by the contract. This outcome will be used to calculate the payment, which will influence contractor behavior. The contract can regulate the behavior of contractor the same way a reference point or set points works in a control system.

A general payment model defines the amount and scheduling of payments the customer should pay the contractor based on the level of effort or outcome obtained from the contractor's effort. In an availability-based contract, the payments are tied to the achieved availability (interval, operational or point) as the result of the sustainment and maintenance activities of the contractor.

The contractor's decision making model assumes that the contractor optimizes their actions to minimize their costs. Using these two players (the contractor and the customer), a two-level optimization can be used to find the contract features and contractor behavior to meet the requirements. One of the important features of availability-based contracts is

23

the time-window that the customer measures the availability over. This concept is called "checkpointing" in some applications and it is the interval in which one measures the interval availability (i.e., the assessment interval).

In some industrial contexts the goal of the contract designer is to optimize (maximize) the availability. For example, the DoD (DoD Directive, 5000) requires program managers (PMs) to "develop and implement performance-based product support strategies that optimize total system availability while minimizing cost and logistics footprint. Sustainment strategies shall include the best use of public and private sector capabilities through government/industry partnering initiatives, in accordance with statutory requirements." However, because the customer cannot know the level of effort of the contract and the contractor cannot be certain about the outcome of their decisions or the customer's level of usage within the time scope of the contract (e.g., asymmetrical information and incentives) it is not a trivial task to enforce and monitor the best effort of the contractor with a contract (Hooper, 2008). For example, when an insured party obtains financial coverage against a bad event from an insurer, they are likely to be less careful in trying to avoid the bad outcome against which they are insured.

Objective Statement

Determine the best outcome-based contract that achieves the availability required by the customer and concurrently minimizes the cost and risk for the customer and maximizes the revenue for the contractor.

The objective addressed in this dissertation can be expressed by the following optimization problem. The contractor is trying to minimize their total cost while guaranteeing a required level of availability over the contract period (assuming N_a equal assessment periods over the fixed length contract).

Customer Goal:

$$\max \sum_{k=1}^{N_a} y_k^* \tag{8}$$

Customer budget requirement:

$$\sum_{k=1}^{N_a} \frac{\text{Payment}\left(y^*\left(\frac{kT}{N_a}\right)\right) - \text{deduction}\left(y^*\left(\frac{kT}{N_a}\right)\right)}{(1+d)^{kT/N_a}} < \text{Budget}\left(\frac{kT}{N_a}\right), \quad k = 1, \ldots, N_a \quad (9)$$

where, $y^*(.)$ is the optimum performance, e.g., availability, d is the effective discount/interest rate per period, T is the contract length in periods and N_a is the number of assessments during the contract time (T). Meanwhile, the customer tries to maximize their profit under the contract requirements (limited by the dynamics of the system),

$$\max \sum_{k=1}^{N_a} \frac{\text{Payment}\left(y^*\left(\frac{kT}{N_a}\right)\right) - \text{deduction}\left(y^*\left(\frac{kT}{N_a}\right)\right) - \text{cost}(k)}{(1+d)^{kT/N_a}} \quad (10)$$

This is constrained by the following financial, performance (availability) and functionality requirements over the whole contract period at all assessment points ($k = 1, \ldots, N_a$).

Financial Requirement:

$$\text{deductions}(k) + \text{costs}(k) < \eta\,\text{payments}(k), \quad k = 1, \ldots, N_a, \quad (11)$$

Where η is the bankruptcy prohibition coefficient to ensure the bankruptcy avoidance for the contractor over the total length of the contract.

Availability Requirement:

$$\begin{cases} y(k) > R_1(k) \\ p\{(y(k) > R_1(k))\} > r_2(k), \quad k = 1, \ldots, N_a \end{cases} \quad (12)$$

where R_1 is a *point availability* requirement as defined in Section 1.3.1 and R_2 is chance-constraint or probabilistic *availability* requirements (refer to Section 1.3.5).

We model the dynamic of the inventory using dynamic system formalism given by

functionality constraints that are derived by dynamical system representation:

$$\begin{cases} x(t+1) = f(x(t), u(t)) + w_1(t) \\ y(t) = h(x(t), u(t)) + w_2(t), \end{cases} \quad t = 1, \dots, T, \qquad (13)$$

in which $x(t)$ is the states of the system, $y(t)$ represents the performance measure, $w_1(t)$ and $w_2(t)$ are models inputs. $f(.)$ and $g(.)$ are functions. It should be noted that we assume the $y(t)$ is completely measurable by customer.

In order to solve the availability-contract design problem posed above, it is necessary to develop a comprehensive and detailed model of information and material flow. The suggested algorithms should provide flexible and robust supply and logistics policies for use in an uncertain environment. The output of such a design activity must be in the form of simple policies, and the performance evaluation should be easily assessed to support the performance of the suggested algorithm or solution. The common parameters in these models address inventory policy (e.g., threshold, lead time, etc.) of safety stock as well as shared inventories and the structure of the supply-chain network. Optimizing the supply-chain network and inventory in a joint scheme is much more beneficial than solving the problems separately. A similar approach should be applied to bring all optimization elements into one platform.

The design process of an availability-based contract will use contract terms, goals and requirements as inputs that define the satisficing parameters for supply chain, inventory management and design parameters of an engineering system with respect to physical-based and budget-based constraints (Figure 5).

The fact that many of parameters might have unknown distributions also needs to be addressed. For defining availability requirements based on the statistics of a fleet of systems, it is not clear what form/distribution of availability is best to guarantee the effectiveness and proper interpretation of the requirement. The role of different requirements distributions is unexplored at this time. There have been many attempts to use algebraic notation to
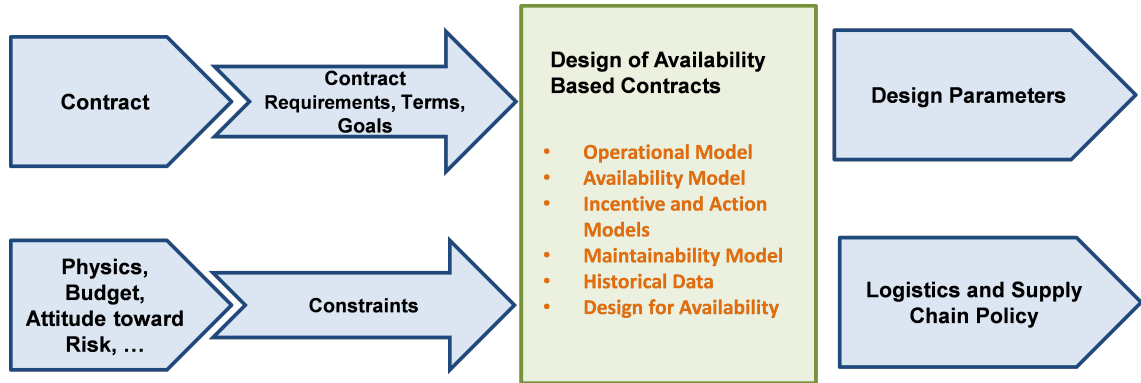
Figure 5: Design of availability-based contracts

formulate the requirement and contract-based relationship, e.g., (Benvenuti et al., 2008).

Approaches to solving this problem should be capable of addressing statistical constraints like the dependence between different variables, there will be no need for over-simplification of the problem; moreover, it should have the flexibility to update the model parameters using Bayesian statistics in response to information added through the life cycle. The difference between assuming fixed end of support contracts and contracts with an uncertain end of support is significant and has been the topic of much of the actuary and medical health-care literature. However, in the maintenance scheduling literature, e.g., (Kim & Park, 2008), end of support is assumed to be fixed despite the common practice of system life extension. The trade-off between cost and availability will be controllable at each point of the life cycle as it is expected to be more expensive to maintain a high level of availability as the system ages. Imperfect maintenance is another inevitable factor that increases cost and uncertainty through the contract term.

2.1   Solution Requirements

Any solution provided should be able to address requirements break-down to lower-levels or sub-contractors. The break down will be used by sub-system designers and they need to benefit from the freedom provided by availability-based contracts otherwise the whole purpose of contract optimization is irrelevant. Solutions are also required to provide a

concise policy for defining logistics policies directly from the availability-based contract requirements. The solution must also be capable of being broken down into contract terms for sub-contractors in terms of bonuses, rewards and penalties. At some point, the degrees of freedom embedded in the availability-based contract will be decomposed and designing a contract should define the best way to do this. Any solution for designing such a contract should consider the uncertainty in the level of incentives for the contractors; and the progressive information gathering that is being done by both customer and contractor.

2.2 Research Tasks

The following tasks have been performed in this dissertation:

1. Gap analysis of availability-based contract design

Extensive litterateur review was done on publication spaces relevant to performance-based logistics and availability-based contracting (details in Appendix B). The requirement and gap analysis was done based on contract theory, defense acquisition and maintenance cost modeling literatures. Then the key findings of this step was used in reviewing modeling, simulation and optimization of reliability and supply chain systems to develop a framework that defines requirements for future solutions as well as elements of the proposed methodology in this dissertation.

2. Model development

Modeling the maintenance process of the system along with the inventory is the base of any further analysis and exploration. The availability and optimal decision making of contractor are taken as key factors of the overall model. A hybrid model that accommodates both the discrete (inventory operation, performance assessment) and continuous (degradation, reliability) nature of the problem has been developed. A discrete-event model captures the maintenance events and reliability features of the underlying system. However, the availability measurement along with the mathematical methods to guarantee optimality of

the method are defined in the continuous space.

The importance of time interval of assessment identified in this stage as well as inter-dependencies in sub-assemblies' failure rate that will cause the number of failures to not be independent in time and therefore a special attention was paid to address this aspect in further steps.

3. Performance measurement design

The goal of availability contracts is to guarantee a specified level of preparedness within a cost constraint. How a contractor is rewarded or penalized is highly dependent on these performance measures. Thus defining the availability requirements is a very important decision customer needs to make. The measurement of these metrics can also be a subject of challenge, due to uncertainties and limited access to measure key performance factors. Moreover, the relationship between these performance metrics and contractor cost model in availability-based contracts is not trivial. Thus, the variety of availability related metrics from variety of real-world contracts collected, classified and compared.

4. Contractor action modeling with an affine controller and convex optimization

The behavior and decision making of contractors modelled within a dynamic programing aimed at cost optimization. An affine controller that represents usage of historical data to make a new decision was used to represent this decision-making. The affine controller also allows the use of a convex optimization scheme and guarantees global optimality in this level.

5. Developing the payment model

A payment model that simulates how different payment structures can be related to the different performance measurement and structure has been developed based on Availability Payments in Public Private Partnership.

6. Optimization of assessment time window for dynamical performance measures

This dissertation proposes the concept of an optimal time-assessment window for which meeting the minimum requirements of a dynamic performance metric will translate into the cost effective and optimum preparedness of the overall project. To find the time-assessment window size, a Monte Carlo analysis was performed on the system and contractor model, and a trade-off analysis of cost and risk is presented to find the optimum time window size.

7. Analysis of the role of uncertainty

The uncertainty of failure rate is different in different systems based on their age or usage rate. Moreover, the impact of uncertainties on contractor's pro-active/dynamic decision making throughout the contract time should be considered. To consider these risks from the viewpoint of customer throughout the contract mean-var analysis used in a multi-objective decision making to isolate contracts with lowest cost and cost-risk.

Chapter 3:   Quantitative Methods for the Design of Availability-Based Contracts: A Review of Methods and Gaps

3.1   Designing Availability-Based Contracts

The relevant literature was classified into several specific groups and the following inclusion and exclusion criteria were used: economic models, operational research models, life-cycle cost models, reliability and maintenance oriented design models, supply chain and logistics surveys and reports. Based on the approach described in the Appendix, the literature is classified. We define contract design as the process of defining requirements and finding the optimum incentives and penalties to impose within the contract. This can include the design of metrics and methods for measuring the desired performance of the contract.

Traditionally, a contract's price and requirements are derived from a life-cycle cost estimation(Bakhshi et al., 2015). Most of these estimations are based on the historical data associated with similar projects in the past, assigning cost to each unit of simulation (Datta & Roy, 2010). Using a variety of information from the past (i.e., reliability, cost, lead-time, delay, etc.), contractors choose the best or optimum parameters for their operation. In this paper, we call parameters pertaining to the sustainment of a system *design policy parameters*. Some of these requirements may be redefined as the project progresses (Defense Acquisition University Press, 2005). The design policy parameters change throughout the life cycle. To find the optimum parameter space for sustainment design to meet the requirements, many simulations over the life cycle of the system are required. Each unique design parameter value requires a model to be assigned to feasible points in

the outcome space (reachability analysis).

In this chapter, we subdivide availability-contract design into its principal elements so that the goals of each portion of the design can be clarified. The following subsections discuss the key elements of this problem.

3.1.1   Reliability Modeling and Condition Monitoring

One of the goals of performance-based contracts is to encourage manufacturers and maintenance (O&M) contractors to improve reliability, aside from focusing on replacement and supply-chain logistics alone. Evidence-based studies have investigated the effectiveness of PBLs on product reliability in the aerospace and electronics industry (Guajardo et al., 2012).

From the view point of the contractor, significant research has been performed on optimizing reliability in the product design stage toward a limited (e.g., for warranty contracts) or unlimited time horizon (Frangopol & Maute, 2003). There is also significant work on multi-objective optimization of reliability and cost along with other performance objectives (Juang et al., 2008; Lapa et al., 2006; Volovoi, 2004b).

From the view point of both the contractor and the customer, maintenance models predict the life-cycle cost associated with different sustainment policies and can optimize the efforts required to maintain a specific level of workload. Bowman and Schmee (2001) outlined the architecture of a simulation tool for pricing maintenance contracts for a fleet of systems using historical data. In the absence of historical data, there are also methods proposed that do not require quantitative records. These methods are well suited to the type of uncertainties in design for long-term development projects as mentioned in (Zietlow, 2007), as well as new product design.

Condition monitoring has been a typical practice for improving availability and reducing maintenance cost. Nilsson and Bertling (2007) have demonstrated the effectiveness of condition monitoring systems in maintenance management of offshore wind farms.

Prognostics and health management (PHM) also can help contractors to meet their availability requirements as demonstrated in (Feldman et al., 2009) and (Lei & Sandborn, 2017). Garza et al. (2008) have introduced the importance of performance sampling procedures for monitoring the contractor's performance under a performance-based maintenance contract specifically from the customer's point of view. Complexities related to contracts that address heterogeneous fleet availability is another venue of research that has been widely ignored. Block et al. (2014) developed a parametric method to measure the fleet-availability of repairable units.

Another factor in contract-oriented design identified in the reliability literature is the stage of life of the platform and the contract time-span. The requirements and conditions for a newly acquired platform are different from the requirements at the end-of-support or at the phase-out stage. Dandotiya et al. (2008) study optimal maintenance decision making for a fleet of airplanes with a variety of ages. Block et al. (2010) optimize the repair scheduling during phase-out of an aircraft fleet by considering the platform end-of-life characteristics.

Human factor reliability and organizational dynamics are also of great importance in the environment of contract-based design. Mendoza and Devlin (2005) demonstrate the importance of organizational design, in maintaining the desired performance in the environment of contract-based design. We will address these factors in the context of the role of incentives in decision-making, in the next section.

3.1.2   *Role of Incentives and Contract Theory*

Contract theory is a well-developed and well-reviewed subject in the fields of law and economics. In the context of service and maintenance contracts, models incorporate incentives (e.g., payments, penalties, rewards) and uncertainties (outcomes of actions, risk taking behavior), and try to identify decisions that lead to optimum outcomes, given asymmetric information on each party in a contract. For example, Jin et al. (2015) used a

principal-agent model to model the contract along with jointly optimizing the maintenance, the spares inventory, and the repair capacity under the game-theoretical framework. They concluded that longer service contracts are preferred by suppliers because they allow the supplier to save on the annualized inventory investment.

The contract terms and requirements can greatly impact the contractor's decisions and potentially reduce the system life-cycle cost and improve the system reliability. (Guajardo et al., 2012) used an evidence-based method to demonstrate the effectiveness of incentives in performance-based contracting on product reliability. Hawker and McMillan (2015) explored the impact of maintenance contract incentives on the energy production of wind farms.

The effects of a contractor's decisions on the life-cycle cost and availability will not be fully known to the customer at the time of acquisition, or even by the end of the contract time period. This effect is generally categorized as a "moral hazard" problem.[1] In an effort to review the limits of effectiveness of performance-based contracting, (Kobren, 2009) noted that this class of contract does not simply shift all the risk to contractors, but also it can add risks (of non-completion/fulfillment and other risks) to the customer side. For example, contractors may choose not to bid for contracts especially in high-risk research and development projects.

The operational research literature contains many papers considering different abstract models and cases in different industries where the government is the sole customer (e.g., healthcare and defense) (Tsay, 1999; Hockley et al., 2011; Kim et al., 2007). Contractors can be classified into two different categories based on how they respond to the risk and incentives: risk-taker and risk-averse. In the context of sustainment contracts, numerous works acknowledge the efficiency of performance-based design concepts without addressing the interaction between contractor and sub-contractors, or considering physical limitations of the system (Scherer, 1964). The development in this field considers a variety

---

[1] "Moral hazard" means that the party that takes the risk will not be responsible for the possible costs and the other party may or may not have information about it.

of contractual mechanism, contractors and customer configurations. For example (Zhu & Fung, 2012) propose modular designs for the interaction between fourth-party and third-party logistics providers in performance-based logistics contracts.

The role of incentives has also been studied in the game theory applications literature. (Ashgarizadeh & Murthy, 2000) introduced a model for the interaction between the customer and the contractor in service contracts. Without considering the effect of incentives on designer decision making, one might not realize the effect of different contract design parameters specially the effect of penalties and awards on the contractor.

In summary, there are many works that use abstract contract models that capture a variety of contractual configurations but there has been no effort to quantitatively integrate the economic and engineering models to demonstrate the impact of these contracts on the designers, the customer and the sustainment process.

3.1.3 *Supply-Chain Management*

Supply-chain management is one of the application spaces that performance-based contracts are impacting most significantly. Contracting in the supply-chain space in an effort to promote efficiency has been studied extensively; however, theoretical work in this field has not found its way into practice (Lafontaine & Slade, 2002). In a meticulous and extensive review of quantitative supply-chain contract design, Tsay (1999) provides a literature review of supply contracts from a modeling perspective and finds that it is not clear what constitutes a contract in the supply-chain contract literature. Most of the works in this space define performance-based logistics as an efficient supply chain with the flexibility to mitigate disruptions and to evolve as necessary (Glas et al., 2013). Elements of the desired solution are being addressed with high-level information management architecture, intelligent hardware allocation/distribution and extensive data collection and monitoring.

There are qualitative works and surveys that address risk management and modeling of PBL applications (Arora et al., 2010); however, there are few works that address a

theoretical grounding or provide empirical studies of this class of contract. In most cases, the life cycle of contracts in these works are short and do not represent the situation of long-term sustainment contracts. The long-term aspect of performance-based contracting is so significant that long-term relationships between the contractor and customers have also been studied as a form of networks of relationships rather than a market of buyers and sellers (Jin & Wu, 2002).

Among numerous works that consider sharing the benefits of efficiency, (Cachon & Lariviere, 2005) show that when forecasts are not credible enough, supply-chain performance falls short of what is expected. Modeling and prediction for variation in demand is a key feature that is the focus of the inventory optimization research. To address uncertainties in long-term contracts, simulation is proposed to account for the stochastic nature of demand and other uncertainties in the environment, such as change of regulations (Komoto et al., 2011). For a concise review of modeling supply-chain contracts see (Tsay, 1999).

3.1.4 Integrated Design and Joint Optimization

For an enterprise-level performance measure like availability, an integrated end-to-end model that includes maintenance, logistics, supply chain and financial cash flow is needed. The benefits of such models have been addressed in recent literature (Godoy et al., 2014; Grossmann, 2012). In designing availability-based contracts, optimizing inventory, supply chain, maintenance, and system design parameters in an integrated scheme gives the designer a degree of freedom that allows true utilization of performance-based contracting. Studying jointly optimal subsystems or integrating different optimization schemes is different from serialization of a set of problems. Similarly, the design for the post-production purchase period is more than just designing a new product and then optimizing the service separately (Baines et al., 2009; Johnstone et al., 2008). The product-service-systems (PSS) literature deals with dynamic interdependencies of product and service in an integrated scheme. The solution needs to be an effective combination of technical and economic

approaches. The first step is to combine the inventory, maintenance, and operational decisions together and form a unified model that provides visibility into the effect of different parameters (Arora et al., 2010; Rodriguez & Vecchietti, 2010). Therefore, the importance of careful integration of logistics, maintenance and supply chain in the design phase is essential.

3.1.5 *Performance Management and Analysis*

The effectiveness of performance-based contracts has been debated (U.S. Government Accountability Office, 2008). Surveys of performance-based contracts show that customers and contractors face serious challenges in defining the terms and conditions of the contract, including the contract's scope, responsibilities, the metrics to be measured, how to measure them, and the translation of measurement to rewards (Gupta et al., 2011). There is uncertainty in what performance analysis metrics need to be addressed. Possible metrics include the time window of the performance assessment, the size of the fleet that is on demand, and metrics to monitor and the weight of each parameter in building an overall availability measure.

It is clear that conventional life-cycle cost methods are failing to address multi-disciplinary product-service-systems (Settanni et al., 2017). Moreover, in the case of availability-based contracts from the customer's view point, availability is a measurable index for the effectiveness of the service provided by the system. From the designer's view point, availability, along with technical performance, cost, and process efficiency are the final goals (outputs) of the logistic and engineering design process. The asymmetry of available information affects the decisions of contractors, while the customer might not be able to evaluate the decision without having a reverse-looking (i.e., historical) model (Datta & Roy, 2010).

Overall performance evaluation will come from the performance assessment of sub-systems. Relating the performance of different subsystems to the performance of the

overall system is also a critical and non-trivial task (Sherbrooke, 1971). One of the reasons that availability and reliability are factors of interest is due to the potential for calculating these performance metrics at a system level solely based on sub-system level performances. The challenge of evaluating inventory management performance is as old as inventory modeling and optimization (Feeney & Sherbrooke, 1966; Sherbrooke, 2006). The time intervals over which performance is measured needs to be chosen very carefully (Ferreira et al., 2009). Cost modeling of availability type contracts will be strongly tied to these performance metrics rather than activities and material flow (Datta & Roy, 2010; Lai et al., 2002).

Performance analysis and metrics of evaluation are the most important factors on the customer side. If requirements are inaccurately defined, performance-based contracts cannot provide the desired outcomes, and both parties will suffer the consequences. Defining performance metrics and evaluating them is embedded in the definition of the requirements and can be viewed as a legal document in the event of disagreements (Goebel et al., 2000).

3.2   Analytics of Methods

Although the pricing of availability-based contracts has been mentioned in a few reports, e.g., (Whitehead & Jagdale, 2008), these reports provide no details clarifying their approach to capturing the complexities and differences of such contracts compared to conventional contracts. As of now, pricing performance-based contracts is largely absent from the academic literature. In this section, we focus on existing methods of modeling applicable to the availability-based contract design problem (these are summarized in the Appendix). More detailed descriptions of these methods can be found in the associated references; thus, only brief overviews are provided here. The bottlenecks associated with using each of these approaches in availability-based contract design are clarified, and the challenges faced by all of these methodologies are summarized in the discussion that follows.

3.2.1 Optimization

The goal of an availability-based contract is to guarantee a specific availability level at all times when the operation of the system is required. This can be achieved by maximizing the availability, however, meeting a minimum availability requirement can also satisfy this goal — the costs and risks associated with these two approaches can be very different. Availability requirements are usually used as a constraint within an optimization problem (Alfredsson, 1997; Dekker, 1996; Hokstad et al., 2005; Immonen & Niemelä, 2008), where the decision parameters can be reliability, preventive maintenance scheduling, system configuration, or supply-chain costs. In most of the existing optimization works, availability is not actually the objective or the control variable. Although many authors address availability, they are indirectly treating it via other parameters related to reliability and maintainability: reliability (McCall, 1965), maintainability (Canfield, 1986) and the spare part supply chain, including: inventory (Alfredsson, 1997), logistics and administrative, etc. (Labadi et al., 2007). Most of the work in this group does not address the contractual requirements over the total operational time.

Trade-off analysis is an essential part of designing such contracts; more maintenance actions will potentially increase reliability at the expense of more downtime. For example, the effect of different inventory policies on the short term and long term costs are different. There are multi-objective schemes that focus on concurrently reducing the costs associated with supporting the system and increasing availability (Taboada et al., 2008).

Overall, high-availability systems have interested researchers from the fields of operational research, management science, computer networks, and reliability engineering (Immonen & Niemelä, 2008; Janakiraman et al., 2004; Sherif & Smith, 2006). The majority of the existing studies represent the availability using analytical expressions (Albright & Soni, 1988; Dekker & Scarf, 1998). There is another group of approaches that look at availability as a state for the whole system (Sato & Trivedi, 2007). This simplification allows the utilization of elegant methods to prove optimality; examples include inventory

optimization using s-S policy (Feeney & Sherbrooke, 1966; Wei et al., 2011).

The existing work proves the feasibility of availability optimization for simplified types of systems; however, when availability needs to be evaluated for large populations of complex systems over the total support time, no single method suffices. At this point, there are a only few examples of existing works associated with maximizing a cost-benefit function that combines the accumulated life-cycle costs associated with a specific system's management (e.g., logistics, maintenance, reliability, etc.) and the availability achieved (Canfield, 1986; Kajal et al., 2013).

3.2.2   System Dynamics

System dynamics models look at the relationship between different factors, for example efficiency, cost and higher-level factors; and drive the dynamics of results by simulating ordinary differential equations between these factors (Angerhofer & Angelides, 2000). Such a meta-level point of view can model the dynamics of the system as well as expert knowledge about the important factors of the model to the study overall performance of the contractor or the system under contract. Classic dynamics of the system addresses system performance over time. For example, degradation of components, evolution of symptoms related to deterioration mechanisms, and the effect of information sharing on maintenance quality, and generally the relation between different causes of performance change. System dynamics uses a network of differential equations and forms decision modeling approaches that are widely used in logistics and supply chain analysis applications (Hussain et al., 2012) as well as in modeling different aspects of public-private partnerships (Angerhofer & Angelides, 2000). In the project management literature, system dynamics is an appropriate tool for modeling the relationship between different decision variables and essential outcomes of interest for the contractor and customers.

System dynamics is generally used in a top-down fashion to simplify the relationships between different elements within a system and time-varying or non-linear parameters, i.e.,

system dynamics solutions usually don't have a view of individual system components. The source of equations that define parameters comes from experts, surveys, and historical data, and may be inaccurate for new projects that lack historical precedent/data. As such, this approach is also not generally flexible for analysis of what-if-cases.

3.2.3 Markov Chains

Markov chains have been extensively used to analyze different aspects of reliability, dependability, risk analysis and in general probabilistic modeling for operation and management. Markov models are the standard framework for prediction of steady-state performance (Caro et al., 2010). They are rooted in modeling different states of operation (e.g., failures, and repair) and ignore the statistical dependency that exists in each component's failure data and its specific maintenance history. Neuman and Bonhomme (1974) address the maintenance policies under which Markov chain models can provide an accurate estimate of reality. Using Markovian methods to model inventory management is well developed in the literature (Albright & Soni, 1988). Markov models and decision trees in general lack the flexibility required for modeling the reality of maintenance and reliability management. Markov chain modeling forces the models states that are continuous as discrete. Examples of such states include maintenance quality, severity of a failure, etc. More discrete states have to be used if one decides to model such states in more detail. At the same time, the uncertainty cannot be well defined, as each component of a system can only be in one state at a time, since each state must be exclusive from other states. Additionally, due to their state-based structure, Markov chains do not provide a clear prediction of the next event time given the time the component has been in a neighbor state (Simpson et al., 2009).

In Markov chain based studies in the reliability and maintenance literature, there is minimal attention paid to the variability and the stochastic nature of each parameter (or lack of knowledge of the distributional properties of parameters) and changes in uncertainty as time progresses. To address this gap, Bayesian extensions of Markov decision processes

suitable for including epistemic uncertainty are introduced as in (Memarzadeh et al., 2014)
Markov decision processes in general do not allow the model to properly include the
knowledge available about the system, which may result in non-optimal strategies. Markov
chains also will not provide a full picture of system behavior under certain strategies. Also
flexibility of decision making and decision parameters during the contract term, in the
performance-based contracts, is mostly neglected by Markov chain approaches.

3.2.4 *Event-Based Simulations*

All the approaches mentioned so far assume that the dynamics of the system are known,
(e.g., in a closed-form) and they analyze the fleet of systems as a whole. The system can also
be modeled using scenario-based methods that simulate each item of the system separately
through different event-paths/sample paths (Fu, 1994). Simulation-based methods consider
components with differing attributes that move from one event to another in time while
including modeling parameters of each component, such as age, maintenance history,
and usage profile. Many recent analyses use simulation for optimization of different
aspects of maintenance scheduling (Wijk et al., 2011). Simulation-based approaches are
especially useful and common when the model grows in size or the integration of multiple
disciplines is required e.g. (Keskin et al., 2010). Monte Carlo sampling is usually used
for sampling from probability distributions of each parameter, as long as one can estimate
reasonable distributions (Marseguerra & Zio, 2000). Karnon (2003) compares discrete-
event simulation and a Markovian process for assessing the effectiveness of health care
policies; due to the importance of each instance of the system, discrete-event simulators
were preferred. Discrete-event simulation tends to offer better representational support for
organizational decision-making processes (Bodner et al., 2009).

Discrete-event simulation (DES) tests different scenarios, along with various behav-
iors of contractors (Angerhofer & Angelides, 2000). Bowman and Schmee (2001) offer a
discrete-event simulation model utilizing historical data of cost and failure analysis results

to evaluate contract price. Ferguson and Sodhi (2011) addresses the role of simulation in performance-based contract design by looking at the PBL contract as a news-vendor optimization problem and advises on the best inventory policies.

Petri nets are a formal discrete-event simulation approach developed for capturing concurrency and synchronization properties. Formal models like Petri models are however more constrained during model development but have a number of advantages over simply writing simulation codes or discrete-event simulators (Volovoi, 2004b). Petri nets can be used to develop models that can easily be verified for deadlocks, conflict of conditions, catastrophic states, and logical errors. Formal methods offer an articulated representation of a system based on mathematical formalism, in which mathematics helps to prove consistency of the specification and requirements while addressing the reliability parameters such as aging of components (Volovoi, 2004a). Petri nets are also used to model multi-party contracts to look for accordance (agreement with no conflict) of the public and private view of contractors, e.g., Aalst et al. (2010), that will guarantee the correct overall implementation of the contract. Meta-heuristic optimizations like genetic algorithms and particle swarm optimization can also be used in combination of a discrete-event simulation as well as being included in the simulation-based category (Kajal et al., 2013).

Lastly, when it comes to cost modeling, event-based simulation is the powerful and flexible cost modeling method. . As Cai and Tyagi (2014) note, most cost models use a combination of historical data or parametric models that are only valid for the conditions under which the data was collected. However, for novel problems such as multi-generation products with complex design phases a new simulation based paradigm is needed. Due to the novelty of performance-based contracts and their impact on performance data, the capabilities of simulation-based methods are needed(Bakhshi & Sandborn, 2017).

Traditionally, the contract and product parameters are defined separately. In recent years, driven by a need for enhancing system reliability, maintainability, and logistics support, attempts to include contract and engineering (performance) parameters simultaneously have been articulated, but have not been done. There are a significant number of papers with a wide array of measures to determine performance, taking both objective and subjective views.

In this Section, the relevant approaches for designing contracts and products are reviewed and the need for a concurrent contract-engineering design is introduced as a key solution to obtain a more realistic overall PSS design.

The correlation between contracts and the PSS design process can be classified into three categories:

1. Engineering/logistics design using fixed contract parameters

In this category, it is assumed that the contract parameters are given as a set of requirements, and they are treated as fixed input parameters in the PSS design (i.e., they are constraints on the PSS design). Hence, the PSS parameters are designed to maximize the operating performance and functionality that satisfies the contract requirements.

Examples of product design processes (hardware and/or software) that include one or more contract parameters, e.g., cost constraints, length of support requirements, etc., are very common. The analysis in (Lei & Sandborn, 2017) is an example of this category of work where PPA requirements (energy price and the annual delivery target are used to perform maintenance planning design for the wind farm). Other examples include (Nowicki et al., 2008) who developed a spare previsioning system to respond to a given performance-based contract from the viewpoint of the contractor. In (Nowicki et al., 2008) the contractor's objective is to maximize profit and the scope of its activity by optimizing the inventory level (the inventory level is considered to be part of the logistics design).

This scheme also includes sensitivity analysis that addresses the reliability of the product.

Less common are PSS design processes that use actual availability requirements. (Jazouli et al., 2014) estimated the required logistics, design, and operation parameters for a specific availability requirement. In this work the developed model connects the requirements on each operational decision regarding repair, replacement and inventory lead-time so that the impact of contract terms can be seen on the logistics decisions. Jin and Wang (2011) studied the impact of reliability and usage uncertainty on planning PBCs incorporating equipment availability, mean-time-to-failure, and mean-time-to-repair.

2. Contract design that uses fixed product parameters

In this category, the contract parameters are optimized for a given PSS. For example, the following contract parameters may be determined: the payment schedules (amount and timing) (Sharma et al., 2010), profit sharing (Hamidi et al., 2014), the length of contract (Deng et al., 2015), the selected contract mechanism (Hong et al., 2016; Nowicki et al., 2008), supply-chain parameters (inventory lead time,[2] back-order penalties, etc.) (Zhu & Fung, 2012), and warranty[3] design could be determined (Arora et al., 2010).

Examples of work in this category include Arora et al. (2010) who studied an integrated inventory and logistics model to minimize the cost of the total cost of supply-chain support (Nowicki et al., 2008) developed a model that designs performance-based contracts with different lengths and contract fees. In this work the contract design is based on a given product with a fixed initial reliability. They explore the opportunity for further investment in improvements in the product's reliability under the proposed PBC to demonstrate a win-win for the customer and contractor through the optimal choice of contact length.

---

[2]The inventory lead time (ILT) was considered to be a logistics parameter determined from an availability requirement. It is also possible that ILT is a contract parameter that is flowed down to sub-contractors.

[3]Although we include warranty design in the list of possible contract design activities that could be driven by the product parameters, for most products that have warranties the type of warranty and its length are determined by marketing, and are not based on the product's predicted reliability. More commonly, the warranty type and length (which are a contract) are passed to the engineering design to determine the appropriate warranty reserve fund.

Hong et al. (2016) employed mechanism design theory[4] to design an optimized maintenance service contract for gas turbines in which uncertainties associated with customer actions, engine performance, and maintenance costs during the contract execution phase were accounted for. They assumed that the gas turbine design was given and determined the contract that maximizes the expected profit and provides a win-win incentive for the customer and contractor.

Wang (2010) developed and discussed three different contract options for maintenance service contracts between a customer and a contractor for a given system design. The contract options were: 1) a full contract that covers both inspections and inspection repairs, and failure repairs, 2) a partial contract that covers inspections and inspection repairs, but not the failure repairs, and 3) a partial contract that covers failure repairs only.

For this category, there are several challenges. The existing models require a better understanding of the impact of incentive structures on the system design and usage. Zhu and Fung (2012) proposed a model based on the service delivery and customer satisfaction level. They studied the design of optimal contracts that balances the incentives and risks to the two sides of a contract, so that both can achieve maximum profits. They assume that incentive payments to the contractor are dependent on the contractor's performance. Further research is also required on the risk attitude of contractors: risk-aggressive, risk-averse, or prudent. In addition, a more general and comprehensive model would include flexibility for the service provider to change their level of effort during the project to increase the chances of meeting their contractual goals. Moreover, an important gap in contract theory models is the assumption of a static risk allocation for the entire length of a project.[5] Zhao and Yin (2011) propose a theoretical model for a dynamic risk allocation in constructing a project. However, a successful dynamic risk allocation needs a comprehensive understanding of both engineering and contractual parameters and their variations throughout a project. Such

---

[4]Mechanism design theory is an economic theory that seeks to determine when a particular strategy or contract mechanism will work efficiently.

[5] This problem is also reflected in choosing a single value for the cost of money, i.e., the cost of money is not constant over time (nor the same for all projects within an organization).

a dynamic risk allocation is not addressed in any theoretical models and is the subject of the next category.

3. Concurrent design of the contract and the PSS

Finally, the concurrent design of both the contract and the PSS would be the ideal solution (for both the customer and contractor) for real applications. However, there are no models that accurately assess and design CfA, dealing with all the risks and uncertainties involved (Rodrigues et al., 2015). One important proposed solution to fill this gap is to use engineering inputs and to find the engineering connections to current theoretical contract models (Hockley et al., 2011). Kashani-Pour et al. (2016) and Alrabghi and Tiwari (2015) reviewed a wide-range existing analytical models in this space and developed a framework for the design of availability-based contracts with consideration of engineering design and incentive structure.

There is an increasing interest in employing PBC concepts to obtain a better mutual understanding between the supplier and the customer. However, the existing literature is primarily focused on solving the problem from the contractor point of view and does not address the role of optimum contract design from the customer's viewpoint. This is partially due to the relatively short history of this class of contract (Rodrigues et al., 2015), a lack of sufficient public data on different design contracts, and ignorance of the dynamic impact of uncertainties in the existing models.

A few authors discuss the need for concurrent design, e.g., (Nowicki et al., 2008) even fewer attempt to provide any type of solution to the problem (Hong et al., 2016), and in cases that claim to address both the customer and contractor, the solutions are primarily sensitivity analyses that ignore the asymmetry of information or moral hazard problem.[6] Another proposed approach (also sensitivity analysis) is to study the impact of engineering parameters on the construction of contracts (Erkoyuncu et al., 2009). Sols et al. (2008)

---

[6] While there are some major manufacturers who appear to (or claim to) use an integrated approach in designing a concurrent contract and product parameters, they are unpublished and no details are available.

studied the formulation of an n-dimensional performance-based reward model for use in PBC contracts. They developed an n-dimensional metrics structure that represents the system effectiveness along with its reward model that results in a successful PBC contract.

The type of cost modeling necessary for concurrent engineering and contract design isn't the same as for either engineering or contract design alone. Most of the current CfA decisions are based on expert opinions, estimation, and historical data from previous designs, which can be unreliable (Knight & Singer, 2014). In addition, such an approach is less useful when system complexity increases (Ferguson & Sodhi, 2011). Also, a lack of relevant historical data is a major source of challenge in new projects (Knight & Singer, 2014; Ferguson & Sodhi, 2011).

Based on Kashani-Pour et al. (2016), solutions provided in this category should be able to address the requirements breakdown (or flow down) to sub-contractors. The breakdown of requirements for use by sub-system designers shares the freedom provided by availability-based contracts. Solutions are also required to provide concise algorithms so that the availability will be tangible and measureable, and so the contractor can implement and understand the requirements within their sustainment activities. Designing availability-based contracts should address reliability design of products and operational decisions based on condition monitoring technologies, the role of incentives and their impact on the life-cycle of the product, supply chain management of the PSS, and the integrated design and joint optimization of different performance metrics. These requirements make the use of concurrent design of PSS and contracts a necessary approach to model the problem for application in real-world practice.

The key questions that should be answered in this category are: 1) What are the main elements of an availability-based contract for a PSS? 2) What are the essential attributes of the concurrent PSS and contract design process? And 3) How are the advantages of concurrent design of PSS and contracts versus the first two category of design verified?

To summarize the concurrent design of contract and PSS needs to address both the

contractor and customer and the dynamics created by the contractual term between them including addressing uncertainties in achieving availability or reliability-related challenges. Concurrent design considers contract design as a part of integrated system design with PSS and the contract of main sub-systems with a dynamic relationship that is subject to stochastic processes such as reliability, supply-chain demand and operational uncertainties.

3.4   Contract Design as a System Design Problem

We approach contract design as a system design problem where the process of designing contractual terms that address performance metrics, the payment model, and performance assessment are design parameters and a multidisciplinary life-cycle simulation of design impacts needs to be integrated into the engineering design process. The significant challenge of contract design in practice is on the customer side.

In the case of availability-based contracts, the TES and engineering designs should determine the contract requirements and the contract length and price in the acquisition and procurement stage, so that it protects the interests of the customer throughout the life cycle (i.e., it does not overpay the contractor, but also minimizes the risk that the system will become unsupported). Also, the solutions provided should be able to address the requirements breakdown (or flow down) to sub-contractors. The breakdown of requirements for use by sub-system designers shares the freedom provided by availability-based contracts. Solutions are also required to provide concise algorithms so that the availability will be tangible and measureable. Hence, the contractor can implement and understand the requirements within their sustainment activities and product life-cycle management.

3.5   Gaps

In order to solve the availability-contract design problem, it is necessary to develop a comprehensive and detailed model that addresses interdependencies, uncertainties as well as the role of contractor incentives. The output of the design activity must be in the form

of straightforward requirements. Figure 6 shows the envisioned design process. The design process iterates, mapping from inputs: contract terms, goals and requirements, to the outputs: satisficing parameters of an engineered system. It starts with the customer deciding on the mechanism of the contract and the best contract-design parameters (e.g., incentive model, contract length, performance objective function). Then the customer needs to model the contractor's maintenance, reliability and supply chain decision making with respect to optimizing the contractor's objective function. This model must consider the uncertainty in the usage and historical data. Since contractor optimization does not address the life cycle of the system outside of the contract, the cost and availability of the system resulting from their decisions needs to be also assessed for its long-term impact on the life-cycle cost of the system. The result will be used by customer to adjust the contractual parameters (during system design iterations) in order to achieve their long-term goals for the system. The constraints on the optimization in each step in the process of designing the contract include both physical-based (technology, geometry, materials, etc.) and budget-based.

The significant challenge of contract design in practice is on the customer side. In the case of availability-based contracts, the design should determine the contract requirements, length and price in the acquisition and procurement stage so that it protects the interests of the customer throughout the life cycle (i.e., so it does not overpay the contractor, but also minimizes that risk that the system will become unsupported). Also, solutions provided should be able to address the requirements breakdown (or flow down) to sub-contractors. The breakdown of requirements for use by sub-system designers shares the freedom provided by availability-based contracts. Solutions are also required to provide concise algorithms so that the outcome will be tangible and measurable, and so the contractor can implement and understand the requirements within their sustainment activities.

Any solution for designing such a contract should consider the variable level of incentives for the contractors (through the term of the contract), along continual information
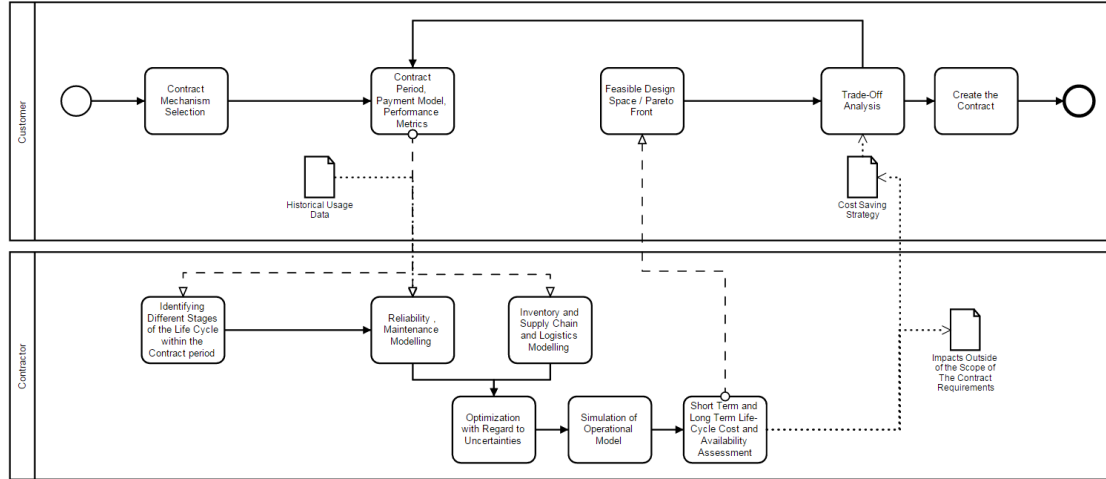
Figure 6: Design process of availability-based contracts

gathering that is done by both the customer and the contractor. Proposed approaches should be capable of addressing statistical constraints like the dependence between different variables. Moreover, a viable approach must have the flexibility to update the model parameters using limited data in response to information added through the life cycle. Optimum length of performance measurement or system condition monitoring for such contracts, as well as the difference between assuming fixed-end-of-support contracts and contracts with an uncertain end-of-support also need further investigation.

3.6 Conclusion

The goal of this chapter was to articulate the quantitative and formal elements of contract models and contract design for sustainment applications in the context of availability-based contracts and evaluate existing methods to address this design problem.

It has been shown that existing solutions are not addressing the degree of freedom provided by this type of mechanism, but they have the essential components of an overall solution. For example, optimizing an inventory will not lead to the optimal availability; however, it is one of the necessary steps in the solution. PBL provides the increased freedom needed to utilize integrated solutions while incorporating the operational risks

involved. Finally, special attention needs to be paid to designing meticulous and effective requirements and performance measures.

By utilizing availability-based contracts, contractors introduce a high-level payment and requirements framework, however bottom-up engineering models addressing the underlying dynamics of the system and the integration of different sub-systems to meet these requirements need to be considered. The feasibility space of contracts and their requirements should be derived by considering the engineering systems with their physical constraints and uncertainties. The integration of engineering design and contract design represents a new paradigm called Contract Engineering. Contract engineering is not a payment structure based on a range of outcomes, rather a careful modeling and simulation of the systems involved is an important component. A Contract Engineer develops a model that can be used for negation by all parties involved and can estimate the impact of different contractual requirements on costs and incentives. Contract Engineering is a practical and engineering approach to guaranteeing a win-win solution space and discovers the feasible regions of design with lower risks for both the contractor and the customer.

This chapter gives program, procurement and acquisition managers' valuable background for assessing the existing cost and decision making models relevant to availability contracting. Using the insight provided, managers can aligned the models and methodologies they are using to availability-based contracting, i.e., determine what models can assess the cost of guaranteed performance considering the integration of all sub-systems involved; understand the operational questions that common methods are not able to answer; can cost saving strategies be compared to business-as-usual practices; and what knowledge do acquisition personnel need to have to assess different cost models, i.e., to perform better negotiation and more accurate pricing?

Chapter 4:   Optimal Performance Assessment Interval Model

One of the key identifying factors of contracts is the role of the assessment time window (also referred to as inspection time window or checkpoints) on the determination of the best possible performance of the contractor. We model the contractor decision making with an affine controller that uses previous data observations to make a decision. This assumption is reasonable for most of the contractor's cost modeling and operational decisions. The affine controller model is then used in a simulation to find the best contracts with optimum assessment window size.

4.1   Background and Literature Review

Finding the optimum checkpoint distances[1] for assessing the condition of a sub-system's performance to ensure the performance of the whole system has been widely investigated in high-availability computer server applications, e.g., (Szentiványi, 2005). Checkpointing helps to switch the system to a backup system, so that the system delivers the maximum availability while not losing its performance. The checkpointing should be short enough to give a small downtime ("failover"), but long enough to utilize most of the system resources for delivering tasks (Szentiványi et al., 2005).

Szentiványi (2005) provides a comprehensive picture of finding the optimal checkpointing distance for high-availability server systems. This paper reviews a group of papers with a variety of modeling detail. All the reviewed methods use queueing theory to model

---

[1]"Checkpoints" are the point that define intervals in which the system performs tasks, intervals are mostly have equal sizes.

the architecture of a server system and then aim to optimize the availability by finding the best checkpoint.

Some researchers investigated the "optimal check pointing interval" problem in the context of fault-tolerant processing systems especially with long-running jobs. Interval availability is availability defined by the amount of time during which the system is in operation over a finite observation period. In this area, there is more focus on the usage and continuous demand. For instance, the amount of crude oil or natural gas to be delivered over a finite period requires related platforms to be available for a certain number of hours within a specified window. Although inventory backups usually cover short interruptions in the production process, the loss of production for several consecutive days might cause problems in meeting the sales contract, involve high penalty costs, and loss of goodwill from customers. In computer and manufacturing systems, the guaranteed performance during a finite period is sometimes a more important competitive factor than the average performance observed over an infinite horizon (Dijkhuizen & Van der Heijden, 1999). In this respect, the interval availability of the production system is often seen as an appropriate performance measure in a practical context; particularly for order-driven manufacturing systems, in which capacity planning plays a key role in satisfying contractual obligations.

Previous work on simulation for contract design has been done by Ferguson and Sodhi (2011) under the assumption of fixed failure rate in which this assessment window was not studied. Ferguson and Sodhi (2011) used a news-vendor in a single-order period model in their simulation-based method to measure the impact of inventory level on the availability of torpedoes under a performance-based contract. Their work explored the secondary metrics that can be used as requirements to help the customer choose a better contractor given their level of inventory. Jazouli and Sandborn (2011) used stochastic simulation by assuming known distributions for logistical parameters to address different aspects of operation over a life cycle. Their direct simulation method determines the design and support parameters that results in a desired availability from the perspective of the
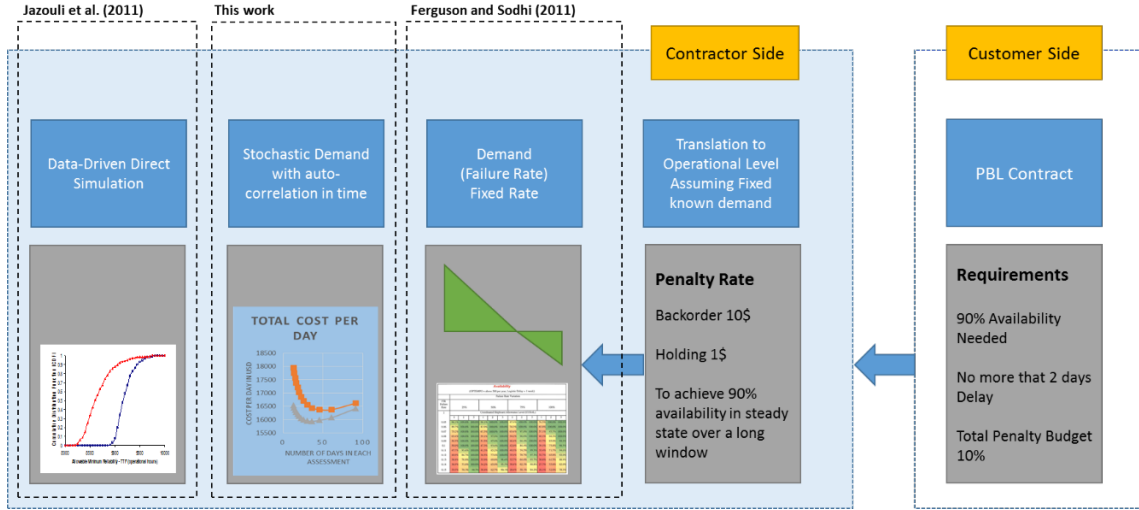
Figure 7: Positioning of the work in this dissertation relative to other key works

contractor. Figure 7 positions the work in this dissertation relative to the work of Jazouli and Sandborn (2011) and Ferguson and Sodhi (2011).

Other researchers, e.g., Schuëller and Jensen (2008), use Markov chain models and consider up and down states for each component and try to formulate the conditions under which a specific number of sub-systems (k out of n) are operational. Faults are assumed to be independent, and subsystems are independent so that there will be a closed-form mathematical representation of the total system availability, which can be optimized to find the best checkpoints.

In the maintenance scheduling literature, the optimum maintenance intervals maintain availability above requirements with minimum cost; however these intervals are not always fixed, and they generally depend on the close-form representation of cost and availability (Kim et al., 2009). Among efforts to address the uncertainty of the demand/failure rate, Verma et al. (2007) use a fuzzy model to find the intervals of preventive maintenance to optimize the cost of maintenance.

From the variability of demand stand point, there are also works in the preventive maintenance space that optimize the scheduling of preventive maintenance to guarantee a level of availability under the assumption of increasing hazard rate using closed form cost

modeling (Kim et al., 2009).

Most of the above existing references formulate a closed-form formula that relates various logistics parameters to availability and cost. Then they assume that demand is independent (in time) and that the contractor has access to infinite resources to support the system (no constraints on the contractor's resources). In existing works that use discrete-event simulators, the contractor is generally ignored and the optimization via trial-and-error without a proof of optimality. Alternatively, to address the problem posed in this dissertation we must consider stochastic demand that is not necessarily independent from time period to time period; and the contractor's behavior in response to incentives must be modeled (i.e., the existing works do not view the problem from a "contract" engineering perspective and therefore are ignorant of the contractor's behavior).

4.2   Model Development

The problem we are looking to solve in this section can be written as

$$N_\alpha^* = \arg\min_{\theta \in \Theta} C(N_\alpha) \quad \text{s.t.} \quad C(N_\alpha) = E(L(\theta, \omega)) \tag{14}$$

In which $C(.)$, the expected total cost of contract from the view point of customer is the performance measure of interest, L will be called *sample* performance, ω represents the stochastic effect of the system, θ is a controllable vector of p parameters, and Θ is the constraint set on θ, defined explicitly or implicitly (by mathematical programing formulations). If $C(.)$ was known explicitly, then analytical techniques including mathematical programing could be usually be applied.

The model for the maintenance operation developed in this chapter uses a convex optimization to design an optimal controller that represents the contractor decision making inspired by (Skaf & Boyd, 2010). The Skaf and Boyd (2010) method uses convex optimization to design a globally optimum affine controller for a discrete-time time-varying linear

dynamic system, perturbed by a process noise, with linear noise corrupted measurement, over a finite horizon. This method addresses the problem of designing a general affine controller in which the control input is affine function of all previous measurements, in order to minimize a convex objective, in either stochastic or worst-case setting. This controller design is not convex in its nature but can be transformed to a convex optimization problem by a nonlinear change of variables that comes below. What follows are the basic steps of such design for a closed-loop controller design.

Considering that the system can be modelled by a discrete-time time-varying linear dynamic system, over time interval $t = 0, \ldots, T$, with dynamics

$$x(t+1) = A(t)x(t) + B(t)u(t) + w(t), \quad t = 0, \ldots, T-1, \tag{15}$$

where $x(t) \in \mathbb{R}^n$ in the system state,

$$y(t) = C(t)x(t) + D(t)u(t) + v(t), \quad t = 0, \ldots, T-1 \tag{16}$$

$$u(t) = \varphi(y(0), \ldots, y(t)) \tag{17}$$

$$= u_0(t) + \sum_{\tau=0}^{t} F(t, \tau)(y(\tau))$$

where we define the feedback matrix as

$$F = \begin{bmatrix} F(0,0) & 0 & \cdots & 0 \\ F(1,0) & F(1,1) & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ F(t-1,0) & F(t-1,1) & \cdots & F(t-1,t-1) \end{bmatrix}$$

which is (m, p) block lower triangular. Then we can have

$$u = Fy + u_0. \tag{18}$$

Then we can solve for x and u in terms of w and v to get,

$$\begin{bmatrix} x \\ u \end{bmatrix} = P \begin{bmatrix} w \\ v \end{bmatrix} + \begin{bmatrix} \tilde{x} \\ \tilde{u} \end{bmatrix}$$

where,

$$P = \begin{bmatrix} P_{xw} & P_{xv} \\ P_{uw} & P_{wv} \end{bmatrix}$$

$$P_{xw} = G + HF(I - CHF)^{-1}CG \tag{19}$$

$$P_{xv} = HF + (I - CHF)^{-1} \tag{20}$$

$$P_{uw} = F + (I - CHF)^{-1}CG \tag{21}$$

$$P_{uv} = F(I - CHF)^{-1} \tag{22}$$

and

$$\tilde{x} = x_0 + Hu_0 + HF(1 - CHF)^{-1}(Cx_0 + CHu_0) \tag{23}$$

$$\tilde{u} = F(1 - CHF)^{-1}(Cx_0 + CHu_0) + u. \tag{24}$$

The matrix CHF is (p, p) block strictly lower triangular, so $(I - CHF)$ is invertible. P is the closed-loop matrix, x as the closed-loop state trajectory, and u is the closed-loop control trajectory. It can be shown that as long as the objective function can be represented as a convex function of P, x, u.

Now that we can use an optimal controller to address the optimal decision making process, there are several approaches that can be used to relate the event-based space to the dynamical representation of the system (time-based).

Integration of an event-based system with a time-synchronous system for simulation can be done in variety of methods and is one of the most pursued goals in simulation

research (Brailsford et al., 2010). It also should be noted that this area of research is not well developed and there are few existing works on synchronization of time-based and event-based methods. The outcome-based orientation of our problem places more emphasis on selecting the integration structure. The proper time-frame to evaluate the performance is one of the key part of this dissertation. Meanwhile the nature of reliability and maintenance actions are generally creating an event-based subsystem (Kashani-Pour et al., 2014).

The goal of the analysis approach is to maintain the preparedness of the system, which translates into insuring a minimum level of availability at all times. For the support of a fielded system this requires management of parts in such a way as to minimize the back-order and holding (inventory position), which will ideally be close to zero after responding to demands in each period. The model involves the integration of the event-based structure (demand generation) with a time-based controller, Figure 8. The time-based controller uses the historical demand data in equal periods of time to determine new order sizes. Demands are generated by a discrete-event simulator that simulates the behavior of the system in time.
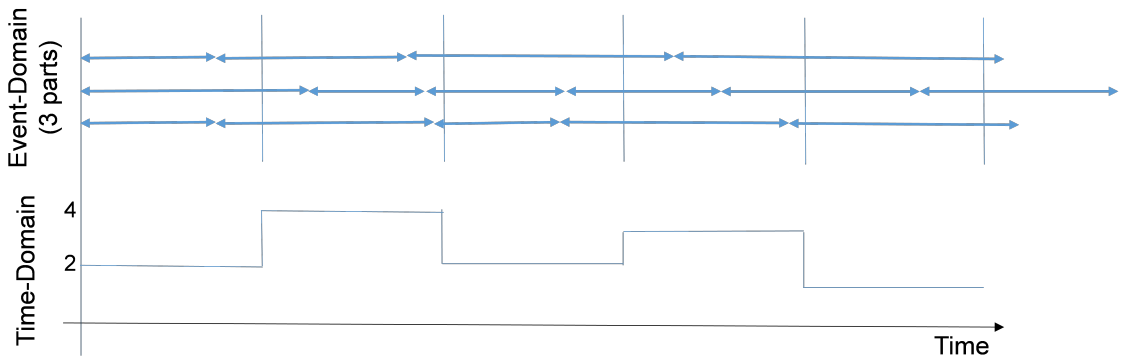

Figure 8: The translation of event-based domain of failure to time-based domain

In (Jin & Wang, 2011) it is shown that product inherent failure rate, usage rate, and the size of the installed base have significant impacts on the equipment availability. Equipment availability is jointly determined by product reliability, usage rate and the size

of installed based which uses these parts. Here the same strategy is used to combine the failure of multiple systems and derive the demand based on the failure of these parts. As Figure 8 shows, the architecture of the analysis approach is based upon a discrete-event retranslation of the process, however the controller only communicates with this model in a time-based regime. Also the performance measurement of the system is a separate activity that considers each simulation path and feeds the controller with a different objective function based on the objective function in a time domain.

The selection of a demand distribution is of great importance. In civil infrastructure (highway management) the demand is selected to represent the condition of pavement or roads, which generally degrades with a slow dynamic, while for operational purposes, systems under PBL contracts consist of parts with a variety of failure rates. Modeling the demand distribution for design purposes has a direct effect on the optimality of the result. In most existing works, demand is considered as being uncorrelated in time, however it seems reasonable to consider a level of correlation in time considering the system level dependencies that these parts might have.

4.3   The Expected Number of Failures

We need to be able to simulate the number of failures in each time interval. For repairable systems, the number of failures at a given operational interval is one of the most important reliability metrics because based on the predicted number of failures, proper resources can be allocated. The most commonly used models for the failure process of a repairable system are renewal processes (RP), including the homogeneous Poisson processes (HPP) and nonhomogeneous Poisson processes (NHPP).

A flexible model (that has been successful in many applications) for the expected number of failures in the first t hours, $M(t)$, is given by,

$$M(t) = at^b, \quad \text{for } a, b > 0. \tag{25}$$

60

$M(t)$ is known as a renewal function.

The repair rate (or ROCOF) for this model is,

$$m(t) = abt^{b-1} = \alpha t^{-\beta}, \quad \text{for } \alpha > 0, \beta < 1, \tag{26}$$

where $m(t)$ is the renewal density function.

The Homogenous Poisson Process (HPP) model has the constant repair rate $m(t) = \lambda$. If we substitute a time-variable $\lambda(t)$ for λ we will have a non-homogenous Poisson process (NHPP) with intensity function $\lambda(t) = m(t) = \alpha t^{-\beta}$. When $\beta = 0$, the model reduces to the HPP constant repair rate model.

Probabilities of a given number of failures for the NHPP model are calculated by a straightforward generalization of the formulas for the HPP. Thus, for any NHPP

$$P(N(T) = k) = \frac{(M(T))^k}{k!} e^{-M(T)} \tag{27}$$

and for the Power Law model:

$$P(N(T) = k) = \frac{[aT^b]^k \, e^{-aT^b}}{k!}. \tag{28}$$

Numerous work have used this assumption and developed models that addresses availability as a function of demand size (reliability). For example (Nowicki et al., 2012) used this model, which is based on the assumption that repair times are independent for calculating the expected back-orders (proxy to materiel availability). It should be noted that NHPP model corresponds to what is called minimal repairs, meaning that the system after repair is only as good as it was immediately before the failure. There are many more possible extensions of such approach which can be derived from statistical modeling literature dedicated to repairable systems (Lindqvist, 2006).

In this work, we create an auto-correlated random values for the number of failures

in time. This time series will be used as the demand stream to test the performance of the contractors under hypothetical contracts.

For the 2-dimensional case: given a correlation ϱ we can generate the first and second values, X_1 and X_2, from the standard normal distribution. If X_3 is a linear combination given by $X_3 = \varrho X_1 + \sqrt{1 - \varrho} X_2$ then

$$Y_1 = \mu_1 + \sigma_1 X_3, \quad Y_2 = \mu_2 + \sigma_2 X_3, \tag{29}$$

so that Y_1 and Y_2 have correlation ϱ.

Likewise for generating n correlated Gaussian random variables $Y \sim N(\mu, \Sigma)$, where $Y = (Y_1, \dots, Y_n)$ is the vector we need to simulate, $\mu = (\mu_1, \dots, \mu_n)$ the vector of means and Σ the given covariance matrix. To use this formula, we simulate a vector of uncorrelated Gaussian random variable, Z. Then we find C such that

$$C C^T = \Sigma. \tag{30}$$

The target vector will be $Y = \mu + CZ$, and a popular choice to calculate C is the Cholesky Decomposition method (Trefethen & Bau III, 1997).

4.4   Discrete-Event Simulation (DES)

Figure 9 shows the general schematic of the closed-loop representation of the proposed system. The availability and several cost factors are chosen as the parameters that the controller needs to control. The demand distribution is derived from the reliability of the parts and the controller action orders new parts for replacement. Control action is defined as an affine function[2] in which we are using previous demands to estimate the new order. Making the control action affine makes comparison of different control policies that can be

---

[2] Affine in the context of nonlinear systems means the control appears linearly (where the nonlinearity with respect to the state is automatically implied).

Contractor decisions are represented as an *Affine Controller*

The *System* represents the model of the maintenance and logistics process

The *Performance* is the outcome of the system in response to a contractor's decision

```
Affine
Controller  →  System  →  Performance
    ↑                          ↓
   Measurement  ←──────────────
```

The *Measurement* is how the payment model in the contract quantifies the contractor's performance for awarding incentives or penalties

Figure 9: Model integration architecture

described by affine functions straightforward, e.g., Model Predictive Controller or Greedy Algorithms. These are common methodologies that use demand forecasting for planning future inventory support. The controller builds a model from a number of samples in the past and then predicts the next demand and the analysis window moves forward in time as more information is gathered.

Because of the complexity and stochastic nature of real world applications, developing mathematical models of the system under study is far from trivial and assessment of their performance is equally difficult. Models that are accurate enough to adequately represent system behavior often cannot be analyzed using, for example, methods based on the theory of continuous-time Markov chains on a finite or countable infinite state space. DES is capable of representing the timeline of the life of different parts and subsystems with fewer restrictions. One can add any number of variables and parameters to the model

without the need to change the structure of model. DES provides a visual indication of what happens to the fleet and each socket. Most importantly, this model provides a probabilistic sensitivity analysis.

DES has the ability to indicate how a supply chain performs and behaves over time when different rules and policies are applied. Testing different scenarios by adjusting parameters and procedures means that supply chain performance and behavior can be explored.

We use a DES model of the platform including its maintenance and we test the controller performance for the system. The parts in this system go from operational to faulty and then based on the availability requirements at any specific time they will be selected for maintenance or replacement. Also, a model of the inventory is provided within the same scheme, and different performance measures can be extracted from this model. Petri nets are a DES approach developed for capturing concurrency and synchronization properties. Petri nets are graphical representations and mathematical tools for formal specification of complex systems (Haas & Shedler, 1986). Formal models like Petri net models have a number of advantages over simply writing simulation codes or DESs. They can be easily and automatically verified for deadlocks, conflict of conditions, catastrophic states, and logical errors in reliability-based design projects (Dohi et al., 2006).

4.5   Controller Mechanism (The Ordering/Planning Strategy)

To model the behavior of contractor with respond to contractual requirements, special attention was paid exploring the decision-making process. A game theoretic two-level optimization problem was used before in Zhu and Fung (2012) with the goal of optimizing the contract from the perspective of customer and contractor separately.

The control-feedback mechanism for availability contracts is based on the established affine control model developed by Skaf and Boyd (2010). As shown in Figures 10 and 11 the model aims to determine the optimal incentives/disincentives in an availability
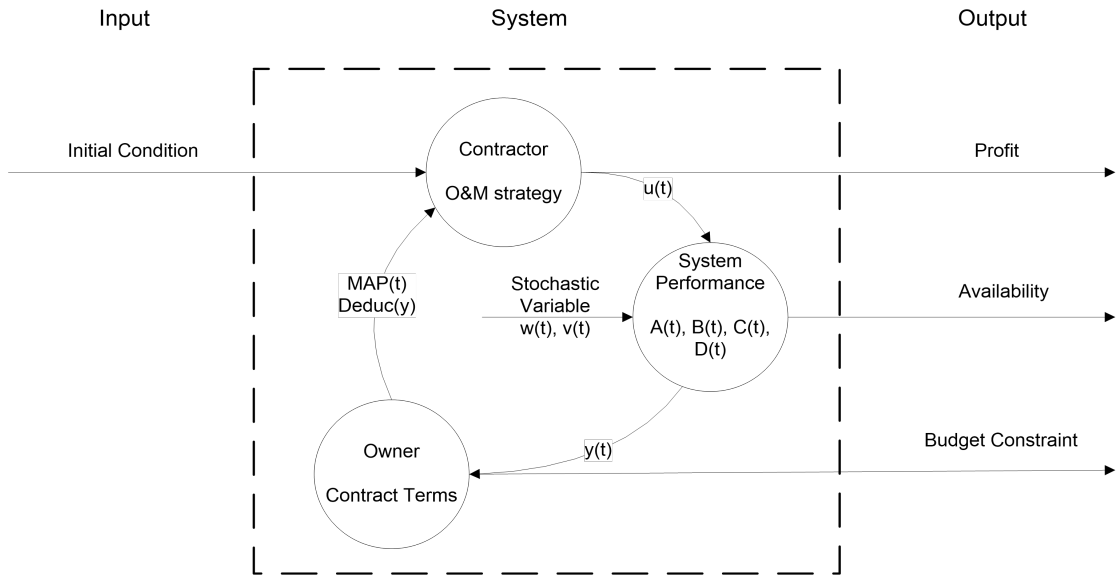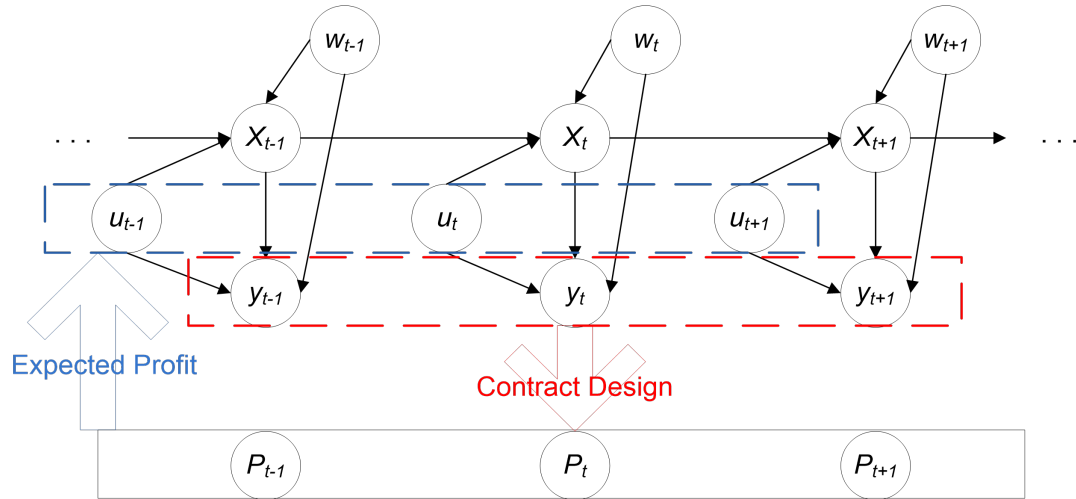
Figure 10: PPP model for availability payment model

X_t System States at time t

u_t Control Strategies at time t

Y_t System Performance at time t

W_t Stochastic Factors at time t

P_t Availability Payment at time t

Figure 11: Affine controller for availability contract

contract so that the customer can expect the best performance or availability given the long-term budget constraint while the contractor maintains a steady revenue (with profit). In public private partnership structure, the private sector, given the MAPs and the deduction matrix, must decide their strategies throughout the operation phase, such as quality of the construction, O&M plan and service quality, so as to maximize their profit and minimize their risk.

4.6 Time Assessment Interval Optimization

The time window that customer uses to evaluation of contractor performance is of great importance. However, most methods target long-term and steady-state performance of simulators. The availability assessment window length is related to T (the contract length) by the following,

$$\text{Avalability Assesment Interval} = \frac{\text{Total Operational Hours in the Contract}}{\text{Number of Assesment During Contract } (N_a)} \quad (31)$$

Also, it needs to be noted that the systems under contract are not operational during the whole contract time and in fact, the contractor needs to be assessed only during the time that the system is operational. For this purpose, we ignore the times that the system is not on-demand and calculate the availability and performance of the contractor based on these operational windows. The assessment also will be performed during one of these operational periods. Figure 12 depicts how the operation time is derived from the total time and how it divides into equal assessment windows. As an example, two cases of quarterly assessment and bimonthly assessment for a 1 operational year contract are shown.

The on-demand time is the time that the customer actually needs the system and the preparedness or availability of system is critical whereas the out of demand times are the time that customer does not require the availability from the system. In out of demand times, the system can be available or unavailable, but it will not count for the hours that the
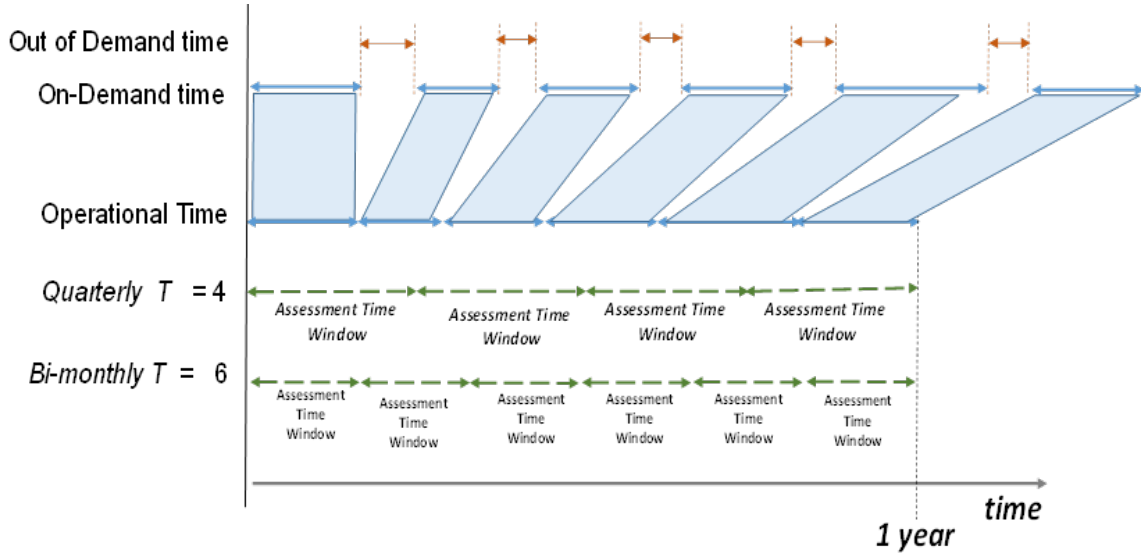
Figure 12: Operational time and time-window assessment

contract requires the contractor to support the system. These are the times that scheduled maintenance can be done without affecting the availability. The contractor's performance is measured in milestones throughout the contract length by assessing the performance only over the operational time, and it will be awarded at the end of each assessment interval. It seems trivial that larger assessment windows (larger assessment window size) will result in fewer assessments (N_a) also, there can be more oscillation in the performance of the contractor. Also, if the time assessment window is too long, then contractor actions near the end of the window will have little impact on the availability measurement (contractors will be inclined to "drop the ball" late in the window because nothing they do will change the result). Alternatively, if windows are too short, contractors are almost penalized for the initial condition of the system and the inventory, and he has a very limited time to learn the demand distribution. Alternatively stated, the size of the assessment window will determine the sensitivity of contractor performance actions to different interruptions and eventually affect the contractor's degree of freedom in design. Optimization of the assessment window size is a primary goal of the model discussed in this chapter. It needs to be noted that in complex systems with high-availability requirements, each availability assessment has an associated customer cost of performing the assessment, i.e., assessing

the contractor's performance is not free. This is another aspect in which contract design is different from optimum warranty design which is usually focused on one agreement period, with no assessment cost to customer (Wu et al., 2007). In Availability based contracts the performance assessment cost can be an administrative cost or the tasks of evaluating the level of availability at the end of each period (option period). For the sake of simplicity, we assumed the same assessment cost for all different length of the assessment window although it can depend on the other parameters (e.g. level of inventory as well).

4.7 Payment Model

Payment model is the second important factor of the contract the shapes the contractor decisions. The following model describes the general form of the optimization process of an availability-based contract:

$$\max \sum_{k=1}^{N_a} y_k^* \quad \text{or} \quad \max \prod_{k=1}^{N_a} y_k^*. \tag{32}$$

In availability-based contracts and models y_k represents the outcome of contractors decisions at assessment number k, which is associated with degradation and reliability model of operating system and it supply chain network. It is also possible to derive y_k without considering the reliability of the system and directly from a dynamical model similar to Sharma et al. (2010). They used linear models to model road deterioration dynamics. In their work since the contracts were long enough that the effect of the transient behavior of the system can be ignored.

4.7.1 Payment Model in PPP Contracts

Depending on the nature of availability and performance measurement the payment model will have different structure. For example in PPP contracts we have developed the below

structure

$$\max \sum_{k=1}^{N_a} y_k^* \quad \text{or} \quad \max \prod_{k=1}^{N_a} y_k^* \tag{33}$$

$$\text{MAP}_k - \text{Deduction}(y_k^*) \le \text{Budget}(k), \quad k = 1, \dots, N_a$$

where y_k^* solves problems ($k = 1, \dots, N_a$),

$$\max \sum_{k=1}^{N_a} \left(\frac{E(\text{MAP}_k) - \text{Deduction}(y_k) - \text{Cost}(k)}{(1+d)^{\frac{kT}{N_a}}} \right) \tag{34}$$

subject to:

$$\text{Deduction}(y_k^*) - \text{Cost}(k) \le \eta \, \text{MAP}_k, \quad k = 1, \dots, N_a,$$

where y_k^* is the availability of the project at the end of k-th assessment interval; d is the effective discount rate per period (more generally the weighted cost of capital); N_a is the number of assessments during the contract time;[3] η is a bankruptcy constraint; and MAP_t (maximum availability payment) and $\text{Deduction}(t)$ are decision variables for contract design for level one (public sector) problem. Given the detailed contract, the private sector (level two) must decide on the best y_k^* at the end of each assessment interval k to optimize its overall profit. To find the best payment plan (MAP) a search on a non-convex feasibility space of the second layer will be needed.

Design space explorations using a variety of search methods and optimization methods is a common approach in multidisciplinary contract-based designs (Nuzzo et al., 2014). In our method, every decision or solution needs to be checked for feasibility of physical system realization. There are variety of methodologies that can be used for this layer from heuristic search to nonlinear-mixed integer programing. Figure 13 shows a possible search method to find the optimum MAP and deduction using a heuristic search.

---

[3]The operational time during the contract is divided into equal independent periods for performance measurement. At the end of each period (time window), an assessment will be done on the level of performance of the contractor. The outcome in each of these periods will determine the payment.
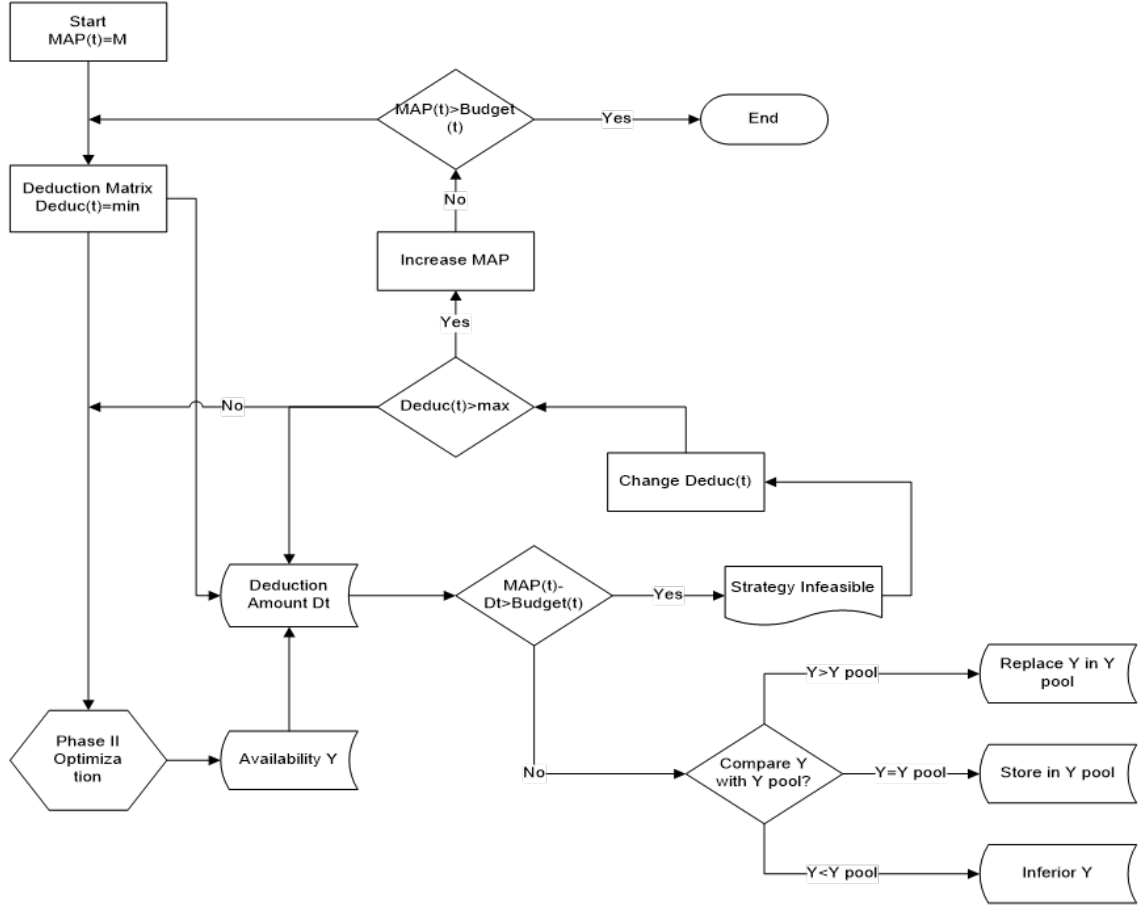
Figure 13: Decision chart of payment model

4.8   The Impact of Penalty Coefficients on Contractor Performance

One of the main tasks of the contract design is to embed the right penalty rates for back-order or holding in the payment model. These penalties are incentives to guarantee the availability requirements while minimizing the total life-cycle cost of support of the system. We can generally assume that there exists a base cost for each back-order of a unit and a holding cost per time unit that the contractor has to pay. Conventionally, when it comes to penalty items, the back-order cost is calculated based on opportunity costs and the cost of downtime. The base holding cost is calculated based on the inventory constraints and inventory operation.

In the process of contract design, the cost model of the contractor is mostly unknown

to customer. But in outcome-based contracts, the customer can only shape the behavior of the contractor with these adding incentives or penalties to these rates at the inventory level.

Chapter 5: Case Study

This chapter presents a case study of a torpedo enterprise. The case study demonstrates that using the modeling developed in Chapter 4, there exist an optimal availability requirements assessment interval for a PBL contract. The case study also explores the best-contracts space from the cost-risk perspective and determines the relative value of using an optimal availability assessment window.

5.1   Torpedo Enterprise System Description

The case presented in this section is based on a case study and data from Ferguson and Sodhi (2011);[1] It examines the inventory of torpedoes for a submarine fleet (Enterprise) managed by a contractor under an availability-based (PBL) contract. The design process presented in this chapter determines the best assessment interval as a contract parameter to reduce the total cost of the system and guarantees that inventory (materiel) availability requirements are met. For our purpose, each item in the inventory is a whole torpedo. Exercises and deployment require constant servicing of the torpedoes, and during testing, if torpedoes are found to be defective, a complete torpedo is replaced from inventory by supply contractors.[2] Following testing, the torpedoes are returned to the fleet for use. When the torpedoes are tested, if they are found to be defective, they are replaced from inventory

---

[1]Note, Fergusoon and Sodhi do not study the assessment window. They used a news-vendor model to measure the impact of inventory level on the availability of torpedoes under a performance-based contract.

[2]In reality, sub-assemblies within a torpedo are tested, however, if any of the sub-assemblies are found to be defective, the whole torpedo is not returned to the contractor/OEM but only the defected parts for repair or replacement. The contractor/OEM guarantees the availability of complete non-defective torpedoes in the inventory.

and the inventory needs to be replenished by supply contractors. The flow of torpedoes in the enterprise is shown in Figure 14.
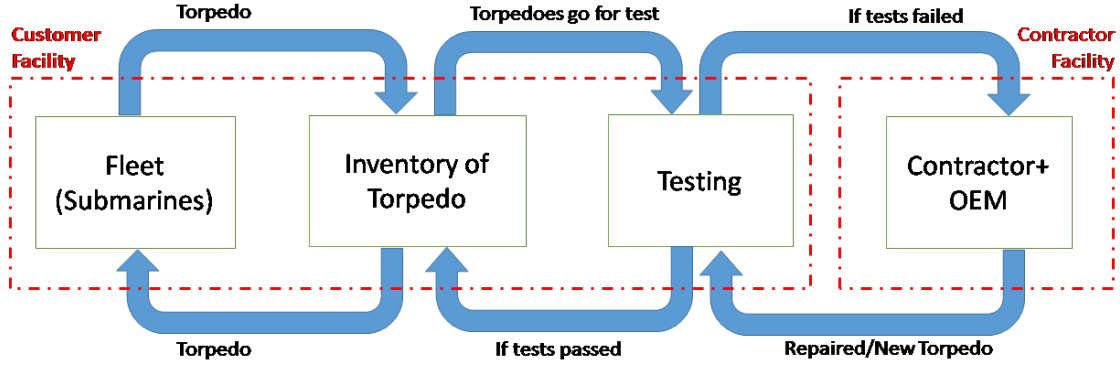
Figure 14: Torpedo Enterprise Material Flow

There are very few works in this application space that report realistic data (Fincher, 2016). We demonstrate our method on this system using the data provided in Ferguson and Sodhi (2011).[3] In their work they look at the torpedo inventory level as an indicator in competitive contracting environment to discriminate between bids.[4] Ferguson and Sodhi (2011) assume that the failure rate of each torpedo is reported as a constant failure rate in a monthly unit. However, in our work we assume a distribution for number of failures with different variance but expected value equal to the fixed rate in Ferguson and Sodhi (2011) to simulate the number of failures over time.

The operational availability defined in (6) in Section 1.3.4 can be used to determine the costs incurred on the contractor side that include shortage, holding, and shipping costs. Back-order (shortage) cost is calculated per day of not having a usable torpedo available in the inventory when one is needed. This could be considered the cost of an inventory (or maintenance) worker's downtime (the work that could have been done), or the cost of penalties due to delays. Storage (holding) cost is the cost of storing one torpedo for one

---

[3]The extent to which the data from Ferguson and Sodhi (2011) represents a realistic torpedo inventory management system is unknown, however it represents a published data set that can be readily used for demonstration purposes. Note, we have added some reasonable data to the case study that was not originally included in the Ferguson and Sodhi work.

[4]The realism of this particular aspect of the torpedo enterprise is unknown, but it is consistent with published cases.

day at the customer's facility.

Table 5.1: Reliability and cost data for the torpedo enterprise from Ferguson and Sodhi (2011)

U.S. Navy Torpedo Enterprise	
	Model A
OPTEMPO	1000 units/year
Total downtime penalty/Unit shipped	$(28)(DTP)
Holding (Storage) Cost	$100/item/month
Back-order Cost	$1000/item/month
Failure rate	$\lambda = 10\% \,$/year
Contract Length (year)	Max 5 Years

DTP: downtime penalty, shipping takes 28 days and its cost is based on the back-order in this case.
OPTEMPO is the expected usage rate of the products being supported by the torpedo inventory for a given time period. In the case of the Torpedo Enterprise, this would be the number of torpedoes expected to be received for maintenance, cleaning, testing and reassembly.

The contract obligates the contractor to support the torpedo enterprise at a specified operational availability while minimizing the number of torpedoes at the customer's facility (the tire example in Section 1.2.1 has a similar requirement). In this PBL environment, the customer/contractor interface is at the "shelf" where the torpedoes are stored at the customer's facility.

The customer provides the contractor with up-to-date inventory levels, which trigger the contractor's decision to restock (replenish spares). For instance, if a torpedo is found to be defective during testing, a spare would be removed from the inventory to replace it.

To model the PBL contractor we use an *affine controller* that makes the order based on the goal of minimizing the cost to the contractor by using the previous demand periods (assessment intervals). The contractor seeks to minimize the number of torpedoes kept in

the customer's inventory by restocking the inventory as the torpedoes are used. In PBL contract the requirements are focused on the final outcome (availability level) and they do not limit the stock level directly and this makes the development of an inventory cost modeling more complex as the costs are not occurring per item and harder to directly connect to a contractor's actions. Table 5.1 shows the data inputs for this case study from Ferguson and Sodhi (2011).[5]

The contract model should address the two different perspectives of customer and contractor. The customer has limited access to the contractor's cost incentives and reliability information and contractor has no role in the customer usage rate of the system. However, from contract design perspective, we are interested in designing a contact from the view point of customer considering including these uncertainties.

The next section describes the details of the stochastic simulation process to generate the number of failures (demand in the inventory model). Section 5.3 explains the basic maintenance inventory model in a general way as the base of the model in this case study. This formalism is not limited torpedo enterprise and can be applied to variety of contract from financial planning, human sourcing to work-order scheduling. Section 5.4 simulates the torpedo enterprise including reviewing the relevant performance metrics and their applicability to the Torpedo Enterprise case. Section 5.6 is the detailed description of optimization-via-simulation for a cost-oriented objective. Section 5.7 considers the risk dimension of the problem and reviews the feasible space from a multi-objective perspective using Mean-Variance (Mean-Var) analysis from modern portfolio theory as a financial risk assessment and optimization.

---

[5]Note, the cost of buying a spare (torpedo) is not included in the model. This is because it is a "wash" between the cases compared, i.e., all cases consume the same number of spares. This case study also assumes that the discount rate can be ignored, i.e., all timeframes are short.

5.2   Stochastic Demand Simulation

In most of the literature related to contract design and reliability optimization regarding performance-based contracts the failure rate is assume to be a fixed (i.e., constant) (Nowicki et al., 2012; Ashgarizadeh & Murthy, 2000). Assuming a fixed failure rate helps produce an elegant closed-form cost model and since most systems spend most of their lifetime in the long flat constant failure rate portion of the "bathtub curve", this is not far from a practical condition in an ideal situation.

In this work we are using two different assumptions for the number of failures in each time interval. First we assume a normal distribution with no co-relation in time and we generate 1000 failure streams for each time period over the contract length. In the second assumption, we look at failures with a lognormal distribution (for generalization over lifetime) and with a correlation in time (to address seasonality, usage rate and common cause failures). For the torpedo enterprise, since the failure rate is reported to be 10% per year we are going to generate different demand streams with the same expected number of failures per year to have a baseline for comparison.

Figure 15a shows five different demand streams (called scenario 1-5) over ten time periods, simulated from the different distribution assumptions. The first scenario is a fixed failure rate; the next four scenarios assume a normal distribution for the number of failures per period (four examples, "sample paths", are shown). Note that we assume corrective maintenance and that's how failure rate and demand are directly related. If the operation is using predictive maintenance, then this number of failures could be sampled from the predicted failure distribution (coming from a Prognostics and Health Monitoring Analysis). In 15b, the five scenarios (five examples, "sample paths", are shown) represent the case where the number of failures is derived from a lognormal distribution with a correlation in time to address seasonality in number of failures.

The results in Figure 15 demonstrate that the assumption of a fixed failure rate (which

Number of Failures Over 10 Time Intervals

(a)

Number of Failures Over 10 Time Intervals - with correlation in time

(b)
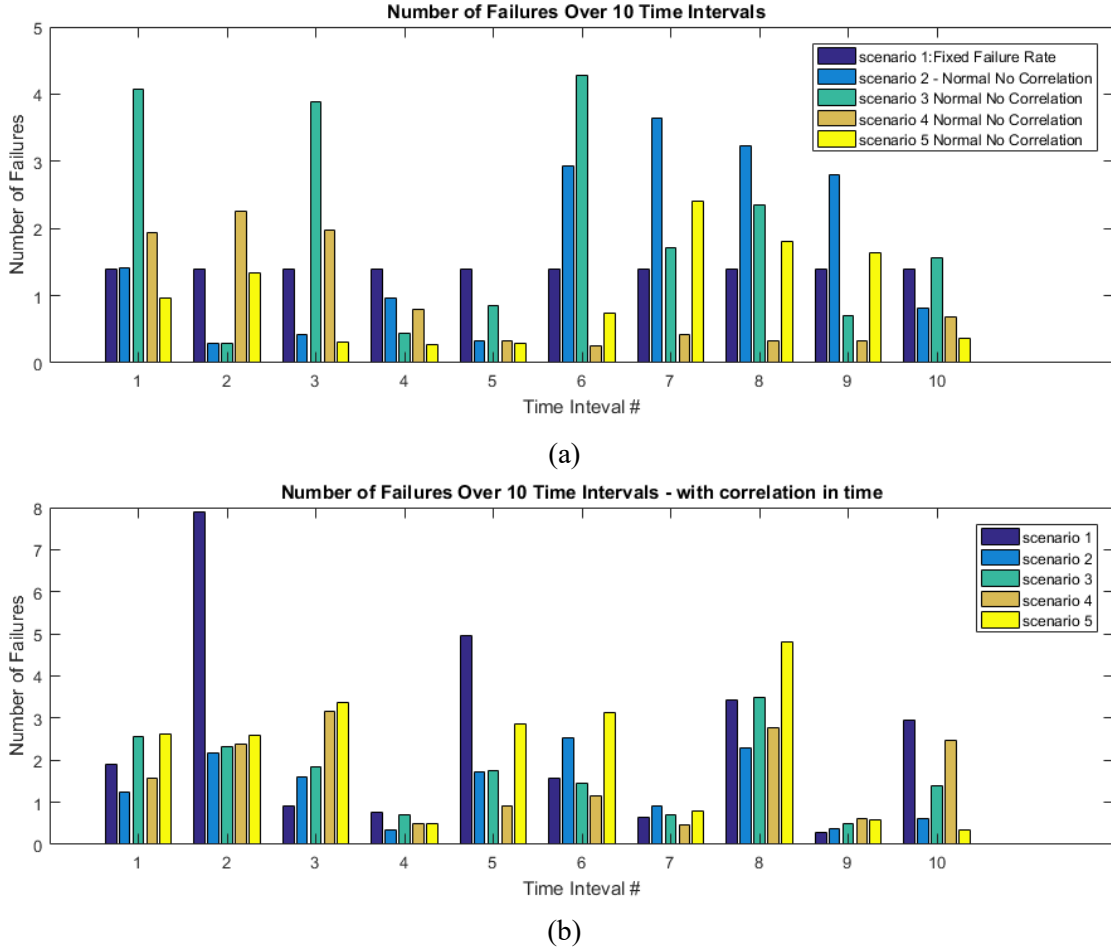
Figure 15: Number of failures simulation over 10 time intervals.

is never really true when time periods are short, due to seasonality and other effects) yields a demand history that varies significantly from a constant.

5.3 Modeling of Maintenance Inventory Replacement

Different maintenance operation and logistics modeling are described in Section 3.2.

5.3.1 Discrete-Event Modeling

For a demonstration of the discrete-event model, consider a repair shop inventory. This is a single commodity linear supply-chain problem. Where we assume that each failure can be replaced or repaired immediately. The decision to repair or replace is being made based

on field standards as well as availability requirements of the spare parts in the inventory.

Figure 16: The inventory & supply chain model developed in stochastic petri-nets

The flow of parts in this system is as follows: a number of systems are running continuously using parts that fail based with a time to failure distribution. At each cycle (e.g., day) after *a* failure events, based on the level of the availability (defined by an operational availability requirement), the part is either replaced or repaired. In this model, we assume a high availability requirement and we do not consider repair, i.e., the model assumes immediate refilling of the inventory with spare parts. We also assume the system will be available immediately after replacement. It should be noted that in reality not all the demand/failures will go for replacement and some of them will go for repair, which will result in longer downtime that could cost the contractor more.

Figure 17: The inventory & supply chain model developed in stochastic petriNets

As an order arrives, based on the replacement request, a new part from inventory is needed. We used stochastic Petri-net formalism (Volovoi, 2006) to develop a preliminary model of this operation and perform preliminary statistical data analysis (Figure 16). Petrinets are commonly used for performance modeling for processes that include stochastic events. Figure 17 shows the graphical representation of the inventory and manufacturing support model, done in CPN Tools (Ratzer et al., 2003). This discrete-event can generate failure data and maintenance inventory demands via sampling from time to failure distributions for the system's part.

5.3.2 System Dynamics / Feedback Systems Modeling

We can formulate the overall performance of this model as a discrete-time linear time-varying feedback system, with the contractor's performance being assessed at times $t \in \left[0, \frac{T}{N_a}, \frac{2T}{N_a}, \dots, T\right]$ and if the performance is not satisfactory the contract will be terminated immediately. At any time the amount of spare parts available to be used in inventory is $(t) \in \mathbb{R}$.

The initial condition is assumed to be $x(0)$. The inventory varies over time defined by the following dynamic:

$$x(t + 1) = x(t) + u(t) + w(t), \quad t = 0, \dots, T, \tag{35}$$

where $w(t)$ presents the demand size (with negative sign) coming from the failure sampling of the system time t in (35). The number of parts available in the inventory at time t can be positive or negative with $x(t) < 0$ meaning a backlog of $x(t)$ units of the parts. The demand for the part at time t is denoted $w(t)$. The number of parts shipped to the inventory at time t is denoted $u(t)$. Figure 18 shows the order of events in a discrete time scope.
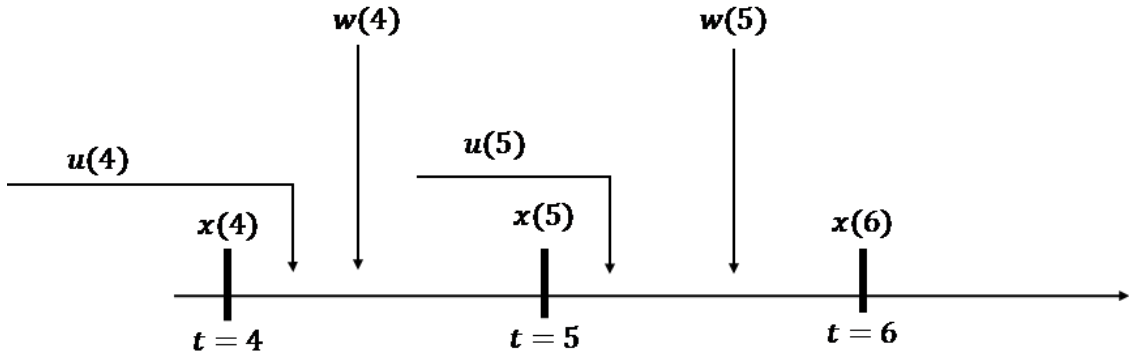

Figure 18: The inventory level at $t + 1$, $x(t + 1)$

5.3.3 Contractor Objective(Cost) Modeling

The goal for the contractor is to maximize profit by minimizing the inventory cost and meeting the customer availability requirements targeting number of units in repair. We assume the contractor addresses the demand in multiple periods during the contract time (as oppose to models based on single-cycle, i.e., Ferguson and Sodhi (2011)). We also assume the contractor addresses requirements by associating penalties to each performance factor as what we call the *inventory cost*.

The inventory cost consists of shipping costs, holding cost and back-order cost. The shipping cost will be proportional to the amount shipped $u(t)$ and inventory costs. For positive inventory, holding cost will be $hx(t)$, and when the inventory is depleted the backlog cost is given by $-bx(t)$. The *total cost* incurred in time t is:

$$\phi(x(t), u(t)) = \max(h(t), -bx(t)) + s|u(t)| \tag{36}$$

and the total inventory cost for the period of $t = [0, 1, ... , T]$ is:

$$C_m(T) = \sum_{k=1}^{N_a} \phi[(x(k), u(k))] \tag{37}$$

where $\phi(x(k), u(k))$ is the *total inventory cost* and the goal of the contractor is assumed to be,

$$\min E\left[\sum_{k=1}^{N_a} \phi(x(k), u(k)).\right] \tag{38}$$

When the demands $w(t)$, $t = 0, ..., T$, are independent, the supply chain optimization problem has a solution of the form of an (s, S) policy (Federgruen & Zheng, 1992). When the demand is correlated across time, and there is no general solution, Skaf and Boyd (2010) use an optimal affine controller approach that performs better than Model Predictive Control and Greedy Algorithm. Moreover, they show that if *demand* has a discrete distribution and can take on only a finite number of values (demand scenarios); the affine controller
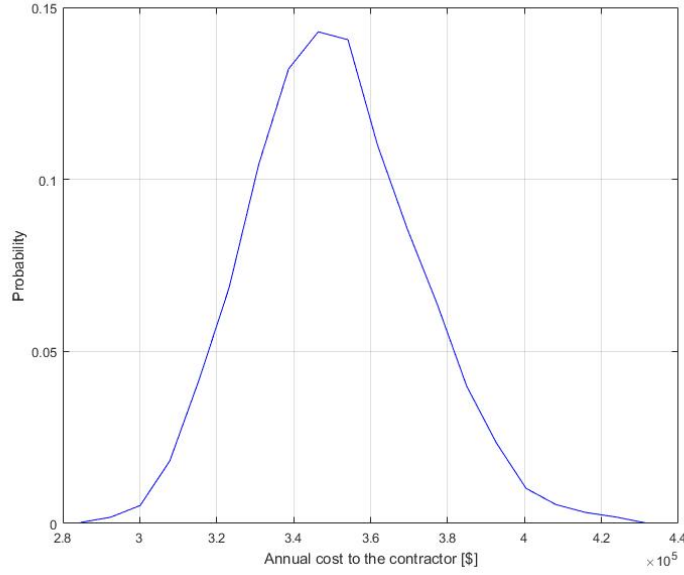
Figure 19: Distribution of contractor costs based on 10,000 simulated contracts

design problem can be reduced to a LP (linear Programming) problem and solved exactly. However, if *demand* has a continuous distribution, the affine problem will have to be approximately sampling from the distribution or by other stochastic optimization methods. To generalize the results, we assumed a real-number demand sampled from a normal and exponential family distribution (Weibull) with different variance to demonstrate and evaluate the suggested model and methodology. The results of simulation of total cost for a contract with $s = 2800$, $b = 1000$, $h = 2000$, $N_a = 10$, and $w(t) = N(10, 2)$ is shown in Figure 19.

5.4   Performance Measurement

We look at the problem of contract design from the view point of the customer. The customer's goals can be described by a variety of *attributes*. System outcomes or functions of the outcomes that we call performance factors that are generally defined by the contract terms, and observable contractor decisions or outcomes of the contractor's actions are the focus of this work. However, there are two types of unobservability that the customer is

facing: 1) the uncertainty of real costs (and their ratios, which defines the contractor's incentives) on the contractor side, and 2) in some cases the customer needs to measure and define secondary functions of these parameters (e.g., operational availability as the ratio of uptime to total operational time, or the ratio of demand to the inventory). Based on the performance factors used by the customer, different measurements and calculations need to be done with the system outputs or the contractor's observable actions (Doerr et al., 2005).

For example, the availability as a function of uptime and the total operational time is a popular measure for operational purposes. For inventory-level contracts (e.g., vendor managed inventories), materiel availability and fill-rate are more common. This is represented in the following equation where $y(t)$ is the availability (α-service level),

$$y(t) = E[q] \text{ such that: } \begin{cases} q = 1 & \text{if } w(k) > x(t) + u(t) \\ q = 0 & \text{otherwise} \end{cases} \tag{39}$$

or

$$y(t) = \frac{\text{Quantity Ordered per period} - \text{Quantity Backordered per period}}{\text{Quantity Ordered per period}} \tag{40}$$

Since this output is not convex based on $u(t)$ we cannot directly use the affine controller method and a non-linear change of parameters or a bisection method is needed.

By measuring the back-order of the different parts and subsystems, we can directly determine the availability of the whole systems and possibly infer the ratio of holding to back-order on the contractor side. This makes the availability the most important factor for measuring the performance of contractors to support complex platforms (Cuthbertson & Piotrowicz, 2011). Other performance measurement metrics are possible.

5.5   Model Setup

Our formulation models the case where the customer makes the decision of assessment interval without the knowledge of the real costs of the contract. As the result, the pricing

decision and selection of performance assessment intervals are made by customer and the contractor will design his strategy around it as a follower. In this section, we lay out the general model that we will use in the game-theoretic (a one shot Stackelberg game formulation) by a two-level optimization. The Stackelberg game is a strategic game in economics in which the leader firm moves first and then the follower firms move sequentially.

In this analysis, we assume the customer information about the failure rate is represented by a random process with the properties described in Section 5.2. We also assume both parties are interested in minimizing their costs.

First, we assume that the total cost to the contractor can be written as:

Total Cost to the Contractor = Baseline Operation Cost + Penalty Costs

The two-level game theoretic model of the problem can be written as below based on previous sections.

Customer Goal:

$$\min_{N_a} C_{customer}(\Phi(N_a)) \tag{41}$$

$$C_{customer}(\Phi(N_a)) = C_{contractor}(\Phi(N_a)) + C_{assessment}(\Phi(N_a))$$

$C_{assessment}$ is based on previous work in assessment interval modeling (Kim et al., 2009), in which there is often a fixed cost associated with each assessment. This model also ignores the money that customer deducts from the contractor payments as they do not help the customer with reduced availability or life-cycle performance-goals.

Contractor Objective:

$$C_{contractor}(\Phi(N_a)) = \max\left(\sum_{k=0}^{N_a}(\text{Payment}_k - \text{Deduction}_k - \text{Cost}_k)\right) \tag{42}$$

$$= \min\left(\sum_{k=0}^{N_a}(\text{deduction}_k + \text{cost}_k)\right) \tag{43}$$

Where the contractor sets their incentive, model based on the customer deduction. We use the model defined in (36) to set the contractor incentive in response to deductions as:

$$C_{contractor}(\Phi(N_a)) = \min_{u(k)} E\left(\sum_{k=0}^{N_a}\varphi(x(k), u(k)) + g(T)\right), \ u(t) > 0 \tag{44}$$

Where $\Phi(N_a)$ is the expected cost of contractor given the number of assessment. The minimization of the contractor also will have a constraint at each inspection period:

$$\sum_{k=1}^{N_a} \frac{\text{Cost}\left(y\left(\frac{kt}{N_a}\right)\right) + \text{Deduction}\left(y\left(\frac{kT}{N_a}\right)\right)}{(1+d)^{\frac{kT}{N_a}}} < \eta\text{Payment}\left(\frac{kT}{N_a}\right) \tag{45}$$

$$0 < u(k) \tag{46}$$

$$\alpha < x(k) \tag{47}$$

The $x(k)$ inventory level at k from (35) and the constraint in (46) makes sure the inventory level stays above the minimum accepted level of inventory during the length of each performance interval. This constraint guarantees that the level of each level will not go below the required minimum performance.

To address the first issue, we assume that there are some baseline values that the contractor is facing, but these values or their ratios are not shared with the customer. This means that the contractor's cost of holding an inventory comes from the actual

cost of running the inventory plus incentives for keeping a lower inventory, given by $h = h_a + h_p$. Similarly, the back-order cost is $b = b_a + b_p$. The h_p and b_p represent the both the incentive/penalty design as part of decision making in the contract design as well as in operational incentives of the contractor. We assume these values are limited by $0 < h_p < 100$ and $0 < b_p < 1,000$.[6] To account for the uncertainty in h_p and b_p, we search the entire feasible space:

$$E(C(N_a)|b,h) = \min_{b_v,h_v} E(\sum_{k=0}^{N_a} \varphi * (x(k), u(k)) + g(T)) \tag{48}$$

So the customer has to choose h_p and b_p as part of decision making in the contract design limited by $0 < h_p < 100$ and $0 < b_p < 1,000$. To account for the uncertainty in h_p and b_p, we search the entire feasible space:

$$h = h_a + h_p, \quad b = b_a + b_p$$

such that

$$\varphi(x(t), u(t)) = \max(hx(t) - bx(t)) + s|u(t)| \tag{49}$$

The total cost of the contract with known parameters ($h = 1000, b = 2000, s = 2800, g = 10,000, N_a = 10$) over 1000 stochastic demand streams (time series of failure numbers) produces the distribution of costs that is show in Figure 20.

Type I(α) availability is the number of times during the contract that the demand was not met perfectly. It does not consider the percentage of the time in which all the failures are replaced immediately with existing stock. Type II(β) is considering what percentage of the failure numbers are replaced, therefore the Type II(β) is generally higher than type I.

The histogram of the Type I and Type II availability and cost shows the average

---

[6]These limits were chosen to simplify the numerical aspects of the simulation. The ratio (10:1) insures that we cover a vast area of design space. Ultimately the important attribute is the points that optimize the contract and the ratio of h_p to b_p plays a bigger role and can be mapped to any other problem including the Torpedo case.
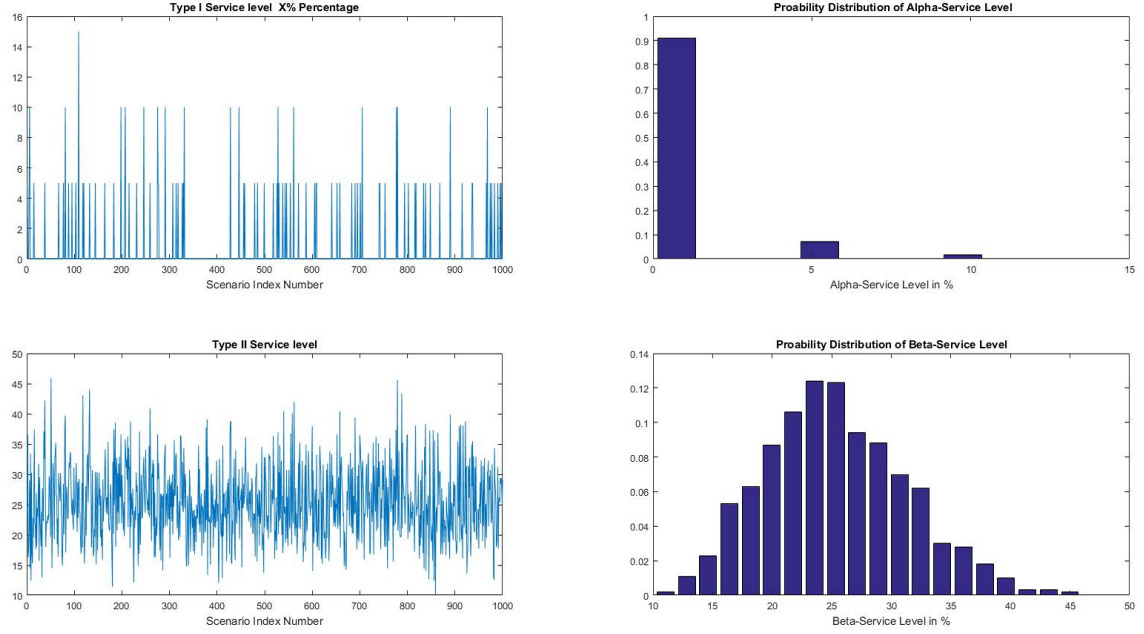
Figure 20: Performance Measurements in a single run for a single contract
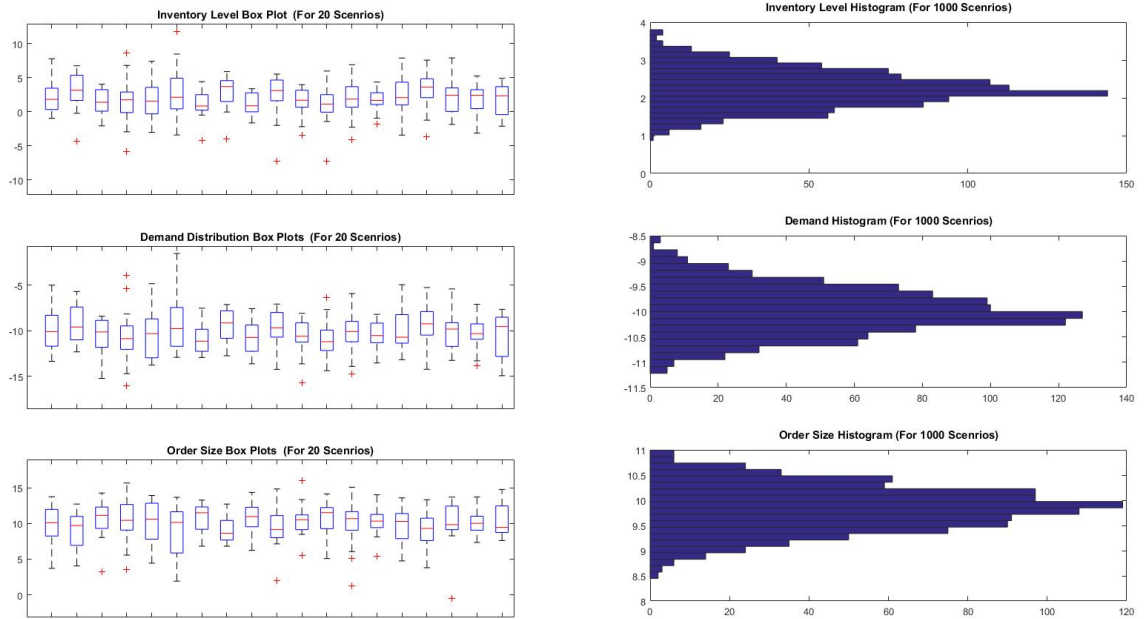
Figure 21: Sample performance scenario in one contract, each performance period is summarized by a box and the histogram is the overall performance of the population over the contract period

number of three performance measures for the given contract over the simulate demand scenarios. Figure 21 looks closer at a single contract with 10 assessment periods and how the inventory level and demand availability oscillates over different demand scenarios. It

is clear that since the contract has a good measure of prediction, the order size in each period meets the demand and the average inventory level at each period is very low. Thus, despite the low alpha and beta service level (materiel availability score), the contractor minimizes the costs with minimum inventory level and back-order size.

5.6   Optimization-via-Simulation

In this section the Optimization-via-Simulation for optimizing stochastic discrete-event systems via simulation is explained. The focus in this work is on the expected cost for each contract using simulation to compare different contract's performance. From the view point of the contractor, there is no need to consider the effect of assessment cost since we assume the assessment is done by the customer or a performance review board on behalf of the customer. To isolate the important factors of contract parameters we also ignore the effective discount rate to make the comparison between different assessment intervals clearer.

Table 5.2 provides the parameters assumed to generate the 900 different contracts that were considered to cover the domain of uncertainty in contractor incentive space from the view point of the customer.

Table 5.2: Parameters of 900 simulated contracts

Holding cost (h) USD/month	Shipping cost (s)	Back-order cost (b) USD/month	Number of Assessments (N_a)	Contract Number
$1000, 2000, ... , 10,000$	Fixed 2800	$1000, 2000, ... , 10,000$	$4, 6, 8, 10 ... , 18, 20$	$1, 2, ... , 900$

The effect of shipping cost was held constant (effectively removing it influence from the analysis) in the model because we are only interested in the behavior of the contractor in terms of reliability- and maintainability-related goals meaning how many items they maintain in inventory to target the availability.

The problem in (48) of Section 5.5 is solved using different methods depending on the particular distribution of $w(t)$. Since $w(t)$ is assumed to be continuously distributed,

we approximated the stochastic problem by sampling from its distribution. Therefore, the expectation in the stochastic objective was replaced by the empirical mean over M samples of $w^{(j)}(t), j = 1, \dots , M$. And the stochastic constraints expanded according to the M samples. Following equations shows the problem in its refined form:

$$\min \frac{1}{M} \sum_{j=1}^{M} \left[E\left(\sum_{k=0}^{N_a} \max(hx^{(j)}(t) - bx^{(j)}(t)) + s|u^{(j)}(t)| \right) + g(T) \right] \tag{50}$$

subject to

$$h = h_a + h_p, \quad b = b_a + b_p$$

$$(hx^{(j)}(k) - bx^{(j)}(k)) + s|u^{(j)}(k)| < \eta \text{Payment}\left(\frac{kT}{N_a} \right)$$

$$j = 1, \dots , M, \ k = 0, \dots , N_a$$

$$0 < u^{(j)}(k), j = 1, \dots , M, \ k = 0, \dots , N_a$$

$$\alpha < x^{(j)}(k), j = 1, \dots , M, \ k = 0, \dots , N_a$$

Where

$$x^{(j)}(t) = (I + HQ)GW^{(j)} + (I + HQ)x_0 + Hr$$

$$u^{(j)}(t) = Q(GW^{(j)} + x_0) + r$$

Due to the sampling procedure, the dimension of the problem increases dramatically, but the problem is still in the form of a linear programming, and remains solvable.

For each contract candidate, 10,000 performance period were generated to test the contractor's optimal decisions over the contract length. A small group of of results are shown in Figure 21 by the box plots of inventory level, failure rate and order size of contractor. The box plot shows the of the performance of the contractor in each assessment interval and its expected value along with outliers caused by large number of contract

population used in the stochastic simulation . The histogram of the inventory level shows that the performance of the contractor is satisfactory not only by the measure of expected value but by the majority of the population of performance. However, because of the large number of samples in the stochastic simulation there are also outliers in performance that are shown in the left plots. In the rest of simulations, in order to illustrate the effectiveness of our framework and model, the following metrics were simulated and studied, expected annual cost for customer and contract along with expected availability measured by service-level metrics.

We also look at the material availability based on the fill-rate definition in (40); Moreover, each point that represents an expected value is addressing the annual performance of a single contract design over 10,000 simulated contractor performance. Considering a fixed cost per assessment independent of the inventory size (assessment is not testing and therefore independent of number of units in inventory or fleet), the contractor performance measured by the expected availability and total cost for a total contract length of one year is simulated, and the result is shown in Figure 22a and 22b.

The calculation of expected annual inventory operation cost from the simulation for 900 contracts is done as follows:

$$E(\text{Contractor Cost}(N_a, b, h)) = \min E\left(\sum_{k=0}^{N_a} \varphi(x_j(k), u_j(k)) + g(N_a)\right) \tag{51}$$

$$= \min \frac{1}{M} \sum_{j=1}^{M} \sum_{k=1}^{N_a} \varphi(x_j(k), u_j(k)) + g(N_a) \tag{52}$$

One can look at the performance of the contractor in each of these contracts for worst-case scenario analysis or derive secondary objectives such as materiel availability as shown in Figure 22a. The points of this plot are coming from the expected $\beta_{service}$ level calculated by:

$$E(\beta_{service}(N_a, b, h)) = E(A|N_a) = \left(\frac{1}{m_b m_h}\right)\left(\frac{1}{m}\right) \sum_{k=1}^{N_a} \sum_{j=1}^{M} \beta_{service} \tag{53}$$
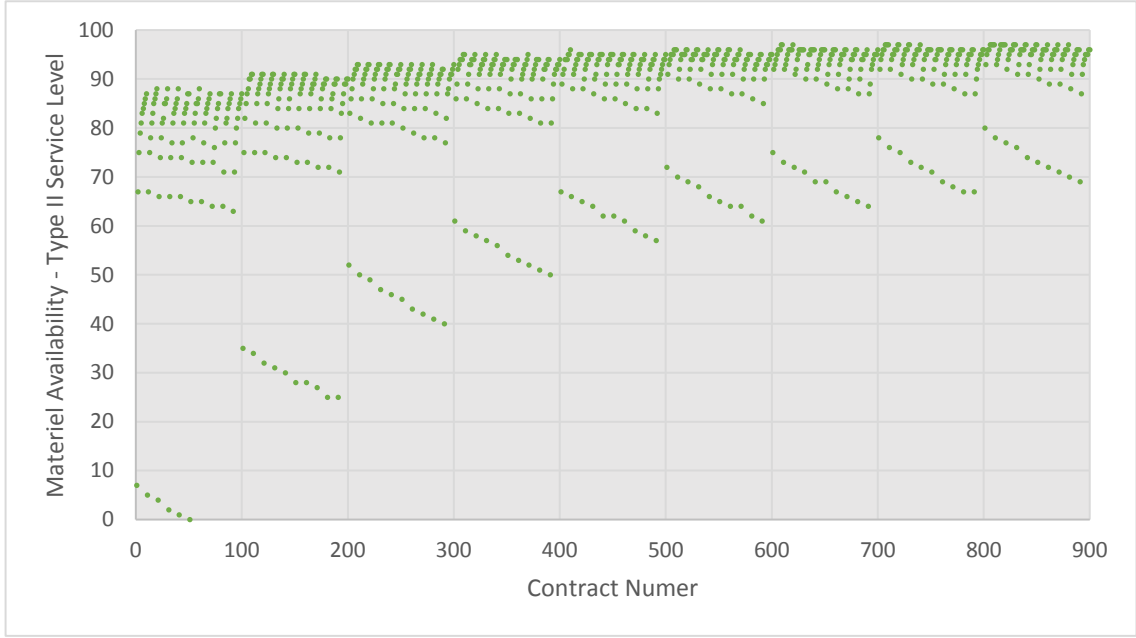
(a) Material Availability for 900 simulated contracts
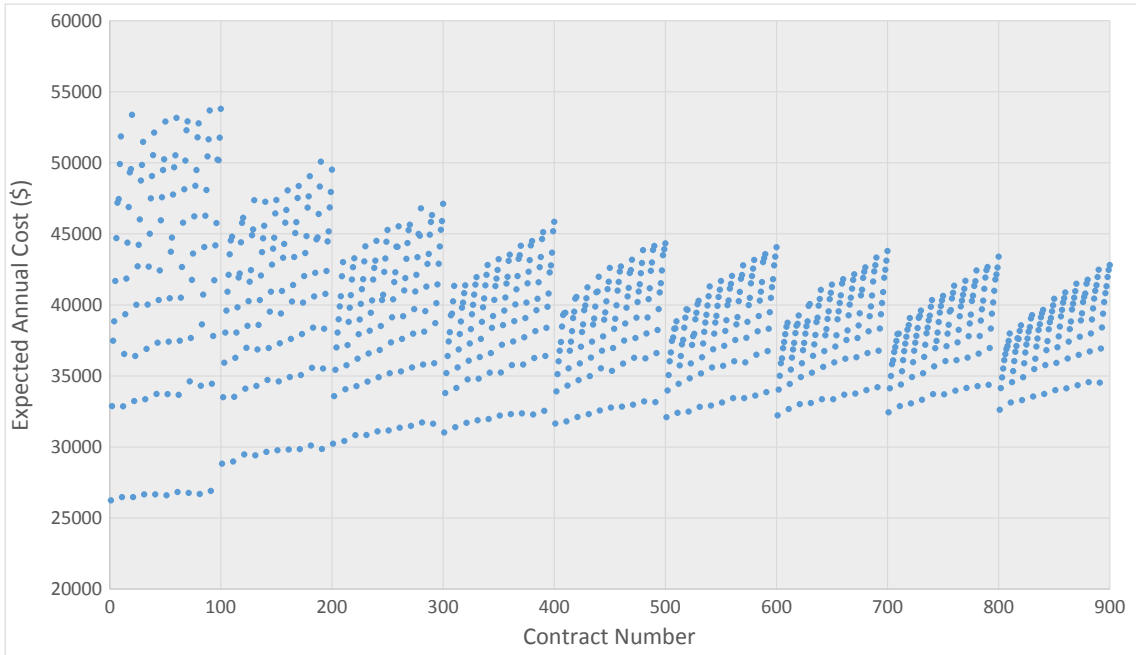
(b) Expected annual inventory operation for 900 contracts

Figure 22: Expected performance measure of contractor in 900 simulated contracts

As Figure 22b shows, the expected cost to the contractor for each of the assessment intervals indicates that by decreasing the number of assessments (N_a), which is the same as increasing N_d, total cost can decrease. However there is a limit to this trend, meaning

Figure 23: Customer Cost including the Assessment Cost

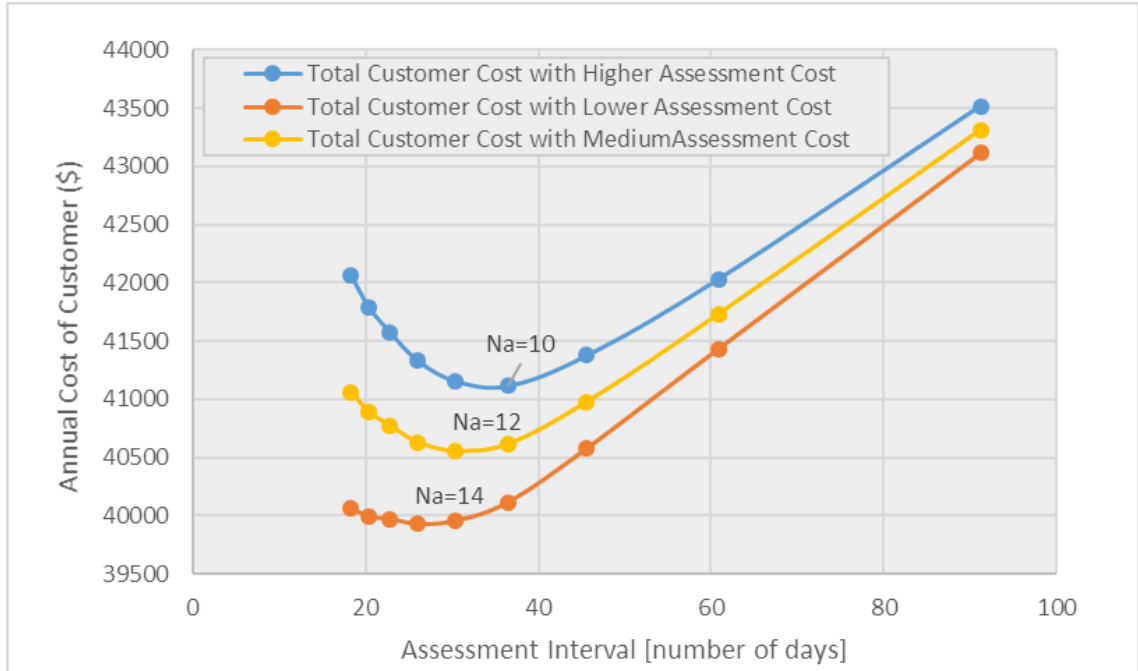that after a certain number of assessments the amount of cost avoidance will diminish. by assuming a cost for the assessments, customer will need a trade-off for the N_a and consequently for the assessment interval. This optimum value of N_a minimizes the cost
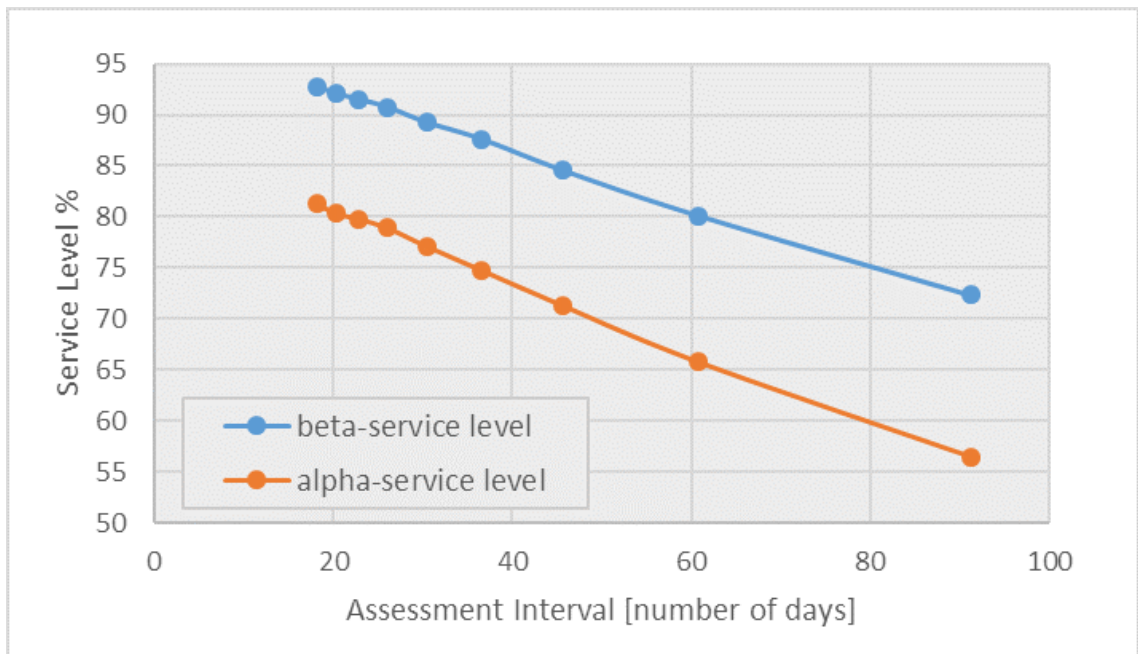
Figure 24: Expected Availability level over number of days in each assessment

of operation for the customer and based on the level of monitoring and payment this can reduce the cost that the contractor incurs as well.

In the next step, we study the simulation results in groups, with each group having same N_a (number of assessments) or N_d (assessment interval length), and use expected value as the descriptive factor of each group, we observe that there is an optimum value of N_d or N_a. Since we are assuming the assessment interval is a customer decision we calculate the performance measures given each assessment interval length. This can be calculated by using all the generated samples. For example customer cost can be calculated by:

$$E(C_c|N_a) = \left(\frac{1}{m_b m_h}\right)\left(\frac{1}{m}\right)\sum_{k=1}^{N_a}\sum_{j=1}^{M}\frac{\text{Customer Cost}}{N_a} \qquad (54)$$

In Figure 23 and 24, each point is a the average performance of 100 contracts that all share the same N_a, and each one is the result of 10,000 sample path simulation for different failure time series. The Service Level is given by (39). Figure 23 is the $E(C_c|N_a)$, represents expected cost for each assessment interval size from the view point of the customer.
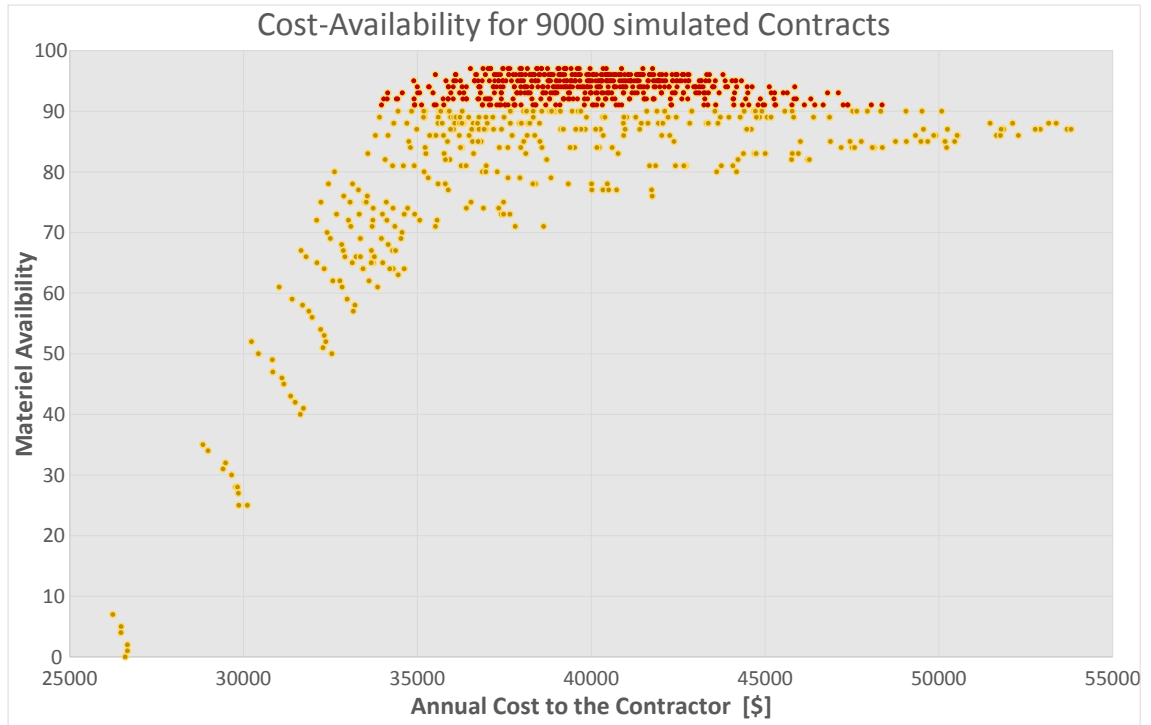
Figure 25: Expected Cost-Availability for 900 Simulated Contract life cycle

The total contract length is assumed to be 1 year. Figure 23 shows that for the customer there is an optimum based on the cost of the assessments. Figure 24 shows that from the availability perspective there is also a shrinking gain from more frequent assessment.

Figure 25 shows the cost-availability relationship in the 9000 simulated contracts. Each contract is represented with their mean of annual cost and mean of material availability as one point on this plot. The group in red are the ones that result in expected availability more than 90% and can be used for further analysis. Grouping the points in Figure 26 by N_a allows the customer to choose an assessment policy, but the customer does not know the costs on the contractor side. Figure 26 shows the direct relationship between cost and availability for each of these groups ($N_a = 12$ has the minimum cost).

Figure 26: Availability-Cost

We have chosen to present the majority of the results in terms of N_a (the number of

assessments in the contract) rather than N_d (the assessment window length) because, N_a more clearly distributes the results in a structured way (i.e., it is linear).

The results show that the optimum interval reduces the cost per period for the contractor as well. The optimal point in the case of low assessment cost ($50 per assessment) as shows in Figure 23, saves between 5-8%[7] of the *total operation cost* of the customer comparing to a quarterly assessment.

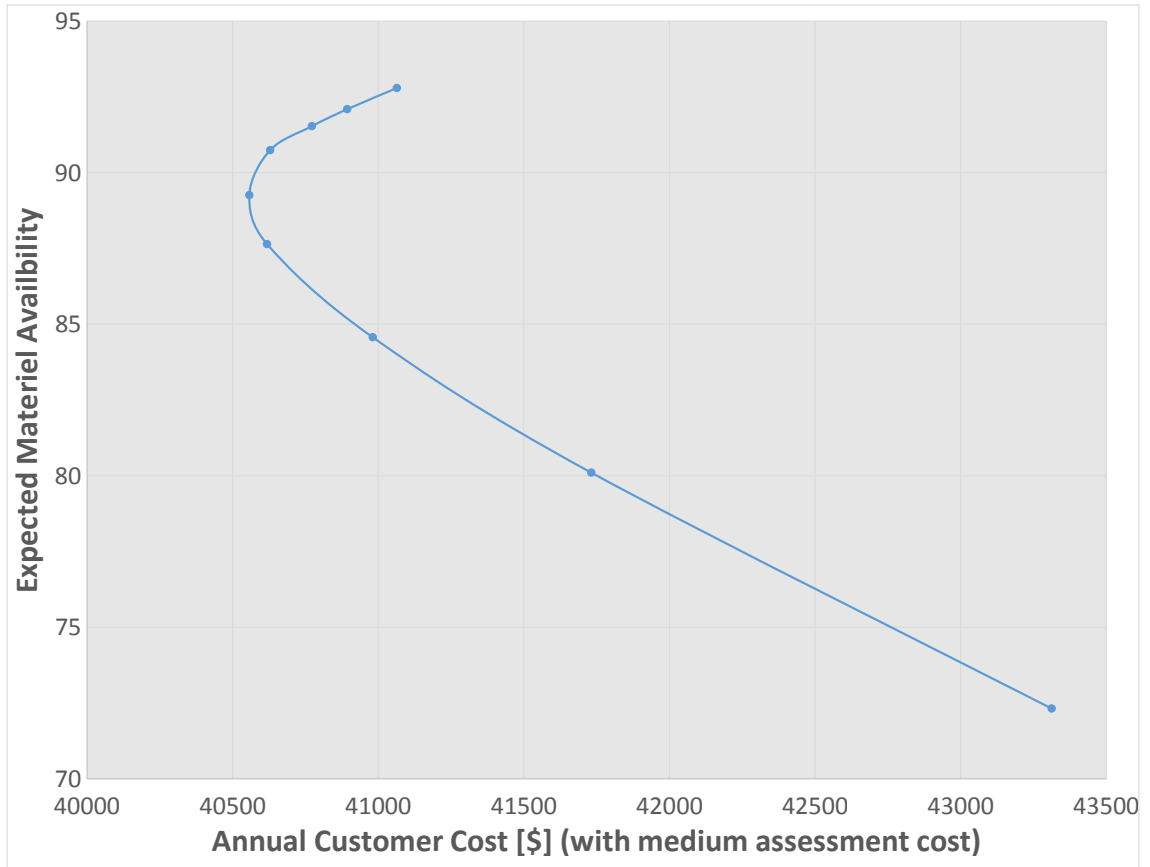Using the proposed affine controller scheme for controlling the availability, we observed: 1) there is a (globally) optimum assessment window length for assessing the contractor. An assessment window that is larger or smaller than this optimum will not benefit the contractor or the customer. However, the cost versus assessment window length relation (Figure 26) is not symmetrical around the optimum point and adding more time to the assessment window has a less detrimental effect than reducing the time (i.e., assessing more often).

5.7   Mean-Variance Analysis

From a risk management perspective, the cost-risk analysis is another critical factor in decision making. Financial risk, in federal acquisitions, is mostly associated with the risk that the project costs more than what was budgeted. The term Cost Risk can be used to refer to this variability of cost from what it is expected to be throughout time. The Cost Risk (CR) can lead to performance risk if cost overruns lead to reductions in scope or quality on the contractor side. Cost Risk (CR) can also lead to schedule risk if the schedule is extended because not enough funds are available to complete the project on time.

A common method that can be used to determine the best contract parameters is the expected variance of cost in each contracting scenario. The variance of costs can

---

[7]In Figure 23, for $N_a = 12$ in the case of medium assessment cost ($100) the annual cost of contract for the customer in a quarterly assessment is $43,312. This means $6.8\% = (43,312 - 40,557)/40,557$ more that optimum cost or in other words the optimum assessments can reduce the cost by $6.3\% = (43,312 - 40,557)/43,312$ from the case of quarterly assessment.

show the range of uncertainty at each decision point for the contractor. This risk will indirectly impact the customer via the contractor performance and readiness. The variance of the expected costs for contractor over 10,000 simulations for each contract is shown in Figure 27. Figure 27 shows that by decreasing the number of days in each assessment, the variability in the performance of the contractor will decrease, which means there will be less variation in cost to the contractor. Figure 27 also shows that if the contractor looks at variance as a determining objective, the previous optimum might not be chosen and depending on the risk-taking attitude of contractor or customer the optimum point can be changed. To find a trade-off between cost-risk and expected cost we use a commonly used method in Modern Portfolio analysis.

Modern portfolio theory (MPT) is a mathematical framework for assembling a portfolio of assets such that the expected return is maximized for a given level of risk, defined as variance. In Modern Potfolio Theory, one models the rate of returns on an investment (asset or decision) as random variables. Here the variance of the expected total cost of a contract is taken as a surrogate for its volatility (risk). Economist Harry Markowitz introduced MPT in a 1952 essay (Markowitz, 1952).

As a measure of risk that contractor is dealing with we looked at coefficient of variance of annual cost for each contract number. In probability theory and statistics, the coefficient of variation (CV), also known as relative standard deviation (RSD), is a standardized measure of the dispersion of a probability distribution or frequency distribution.

$$\text{Coefficient of Variation } CV = \frac{\sigma}{\mu} \tag{55}$$

$$\sigma = \sqrt{\frac{\sum (\text{total\_cost}_{(j)} - \mu_j)^2}{n - 1}}, \quad \mu = \frac{\sum_1^j \text{total\_cost}_{(j)}}{M}$$

As Figure 28 summarizes, a risk-averse decision maker will prefer a situation with larger number of assessments (N_a), which corresponds to a smaller assessment window, to
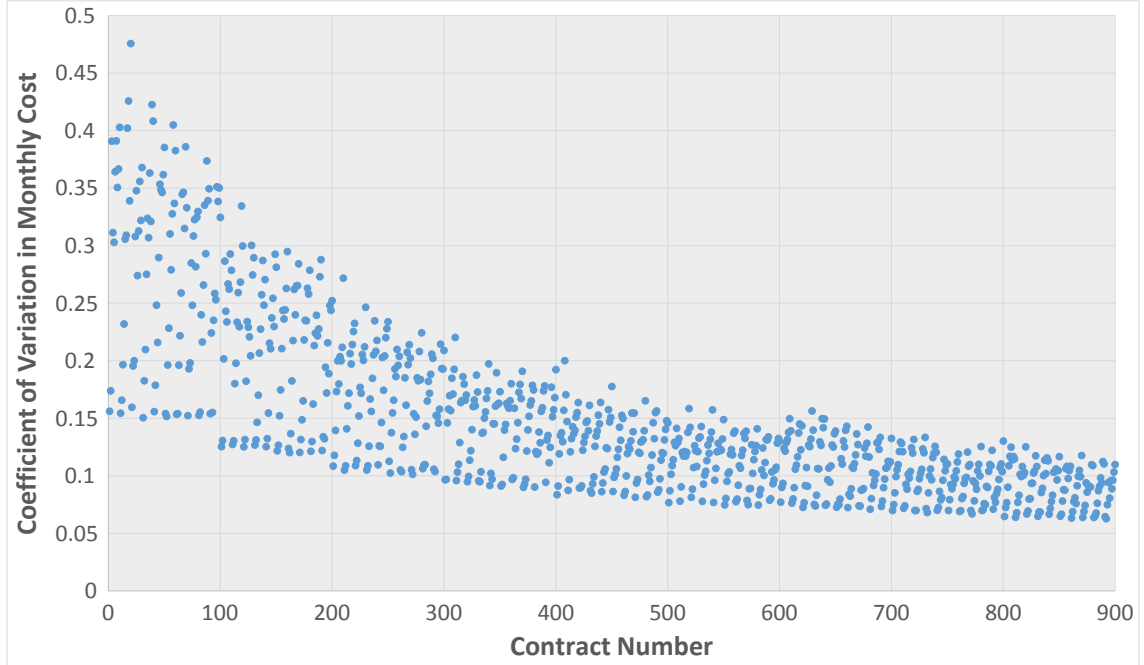
Figure 27: Coefficient of variation for simulated contract lifecycles as a measure of cost-risk for contractor

achieve a lower variation in costs lower due to its lower variability.

It is often expressed as a percentage, and is defined as the ratio of the standard deviation and the mean. Complementary studies such as Value at Risk (VAR) can benefit from the distribution of each simulation.

Optimizing the expected cost alone cannot guarantee that the realized cost measure will fall within a narrow range of its expected value when the corresponding variance in failure and eventually costs at each assessment point is high. Moreover, just focusing on the expected cost ignores the risk attitude (risk aversion, risk neutral or risk seeking) of contractors, which may cause them to change their strategy during the contract period. To investigate this case, we have carried out a mean–variance (Mean-Var) analysis on the simulated contract with their demand scenarios.

$$E(C_c|N_a) = \frac{1}{mN_a} \sum_{j=1}^{M} \text{contractor\_cost}_{(j)}$$

To measure this cost risk, we look at the coefficient of variance of the cost over

97
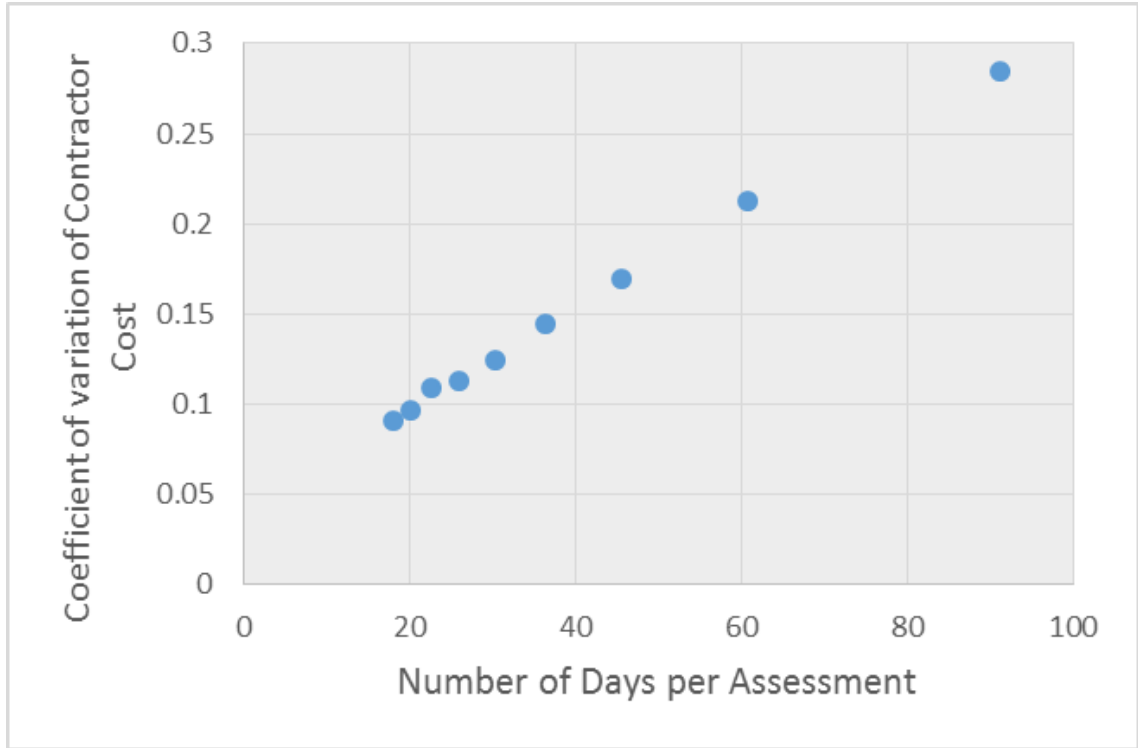
Figure 28: Coefficient of variation for simulated contract for

the contract length. Figure 29 shows expected cost versus the expected risk in 10,000 simulated contracts. Mean-Var Analysis (MVA) recommends that the best contracts will be on the Pareto-frontier of the curve that bounds these data points. The best contracts can be selected and compared against the worst contracts over the same set of demand scenarios and further analysis can be performed.

Since the assessment cost is added linearly it will only shift the mean of the costs and it will not impact the variance of the costs. Therefore, the variability of the assessment cost will not influence this result.

The result of MVA is shown in Figure 29, the hyperbola-shape area shown by the data points is referred to as the 'Markowitz Bullet'. The efficient hyperbola (frontier) of this curve is where the risk is minimized and the return (cost savings in this case) is maximized.

We can trace back the points on the frontier and find the contracts that produce the frontier of this group. Figure 30 uses the same group average to show the effect of the
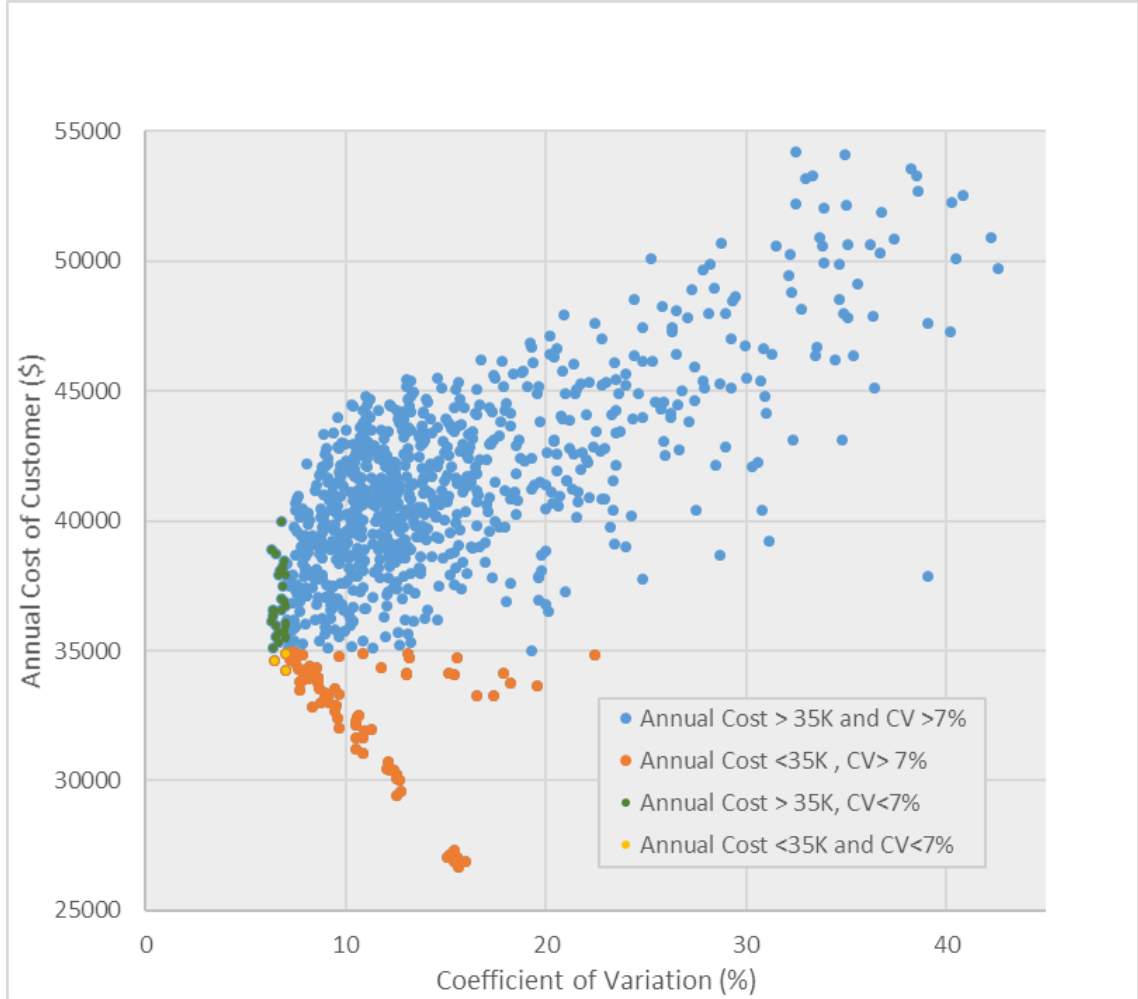
Figure 29: Mean-Var Analysis of 9000 contract from the view point of contractor

assessment interval on the cost-risk curve for the customer. The lowest cost and lowest risk is highly desired by the customer.

5.8   Discussion

In this case study, a stochastic model is used in a hybrid simulation to search for the optimal number of assessments throughout an availability-based contract.

A non-collaborative relationship between contractor and customer is assumed in the sense that contractor does not share information (cost structure, reliability) or profit with the customer. Since we are looking to design a contract from the customer point

Figure 30: The Mean-Var analysis from the customer point of view

of view assuming an *ideal* contractor, we first assumed a range of uncertainty for the operational cost of a contractor given different contractual penalty rates and assessment intervals. We also assumed that the contractor will try to maximize its profit by minimizing the spare parts inventory operational cost while respecting availability requirements. We used an affine controller to model the contractor behavior, which uses historical data to make an optimum decision in every ordering cycle. This algorithm has been shown in previous work (Skaf & Boyd, 2010) to outperform Model Predictive Controller (MPC) and greedy algorithms. Using these assumptions, a hybrid of discrete-event simulation and dynamic system simulation was used to test 900 contracts each with 1000 failure time series. Unknown information about contractor was addressed by assuming uniform distributions. The customer will ask for a certain number of assessments (N_a) throughout the contract. The cost of one year of the contract was used as a basis for comparison between different contracts. Data provided by Ferguson and Sodhi (2011) provided us with a practical range

of the inventory parameters for a Torpedo Enterprise. The cost to the customer will also include the cost of performing the assessments. Next, the cost-risk issues were addressed by employing Mean-Var analysis. The coefficient of variation (CV%) of cost per period was used as the measure of risk that contractor and customer are facing at each period. Finding the best contract based on the cost-risk criteria is a multi-objective decision that uses Mean-Var analysis data.

In the formulation of affine controller for the contractor, we used a convex cost structure to achieve a global optimum, which minimize the inventory level and back-order. This model can be extended to target a certain level of inventory (safety stock), order-size limits, or an availability growth curve through time.

An availability-cost trade-off analysis was used to show the impact of a different number of assessments including its associated costs. In our analysis, we assumed that reliability (and thereby demand) is not a control parameter for the maintenance contractor, however in reality, PBL is designed to incentivize OEMs to improve their reliability (Guajardo et al., 2012).

To achieve a real-world application impact, it is assumed that the number of failures in each time per period are not independent and are correlated to address the imperfect maintenance, the seasonality of failures and other factors identified in field-data reliability literature (Yigitbasi et al., 2010). The performance assessment and payment structure in this work was based on extending a PPP modeling.

Using this stochastic model, it is shown that when there is a complete lack of information about contractor's incentive and cost structure and without requiring detailed requirements, the customer can build a framework by assuming an ideal contractor and the customer can still design a contract that reduces the total life-cycle costs.

The model aims to provide guidance for better design and negotiation of availability contracts and is expected to help both parties understand the essential purpose of the partnership, and seek their mutual interest more efficiently. The optimum number of

assessments and time interval for assessing the contractor's efforts is a key factor that determines the constraints for a contractor's design process. Longer assessment interval allows for more fluctuation in the inventory level, but the prediction of demand can be done more accurately by contractor. The length of the assessment window translates directly to the length of time over which availability is measured for contract assessment purposes.

An end-to-end quantitative model of operation that addresses the contract parameters (assessment interval, penalty and payment model) supports contract design negotiations and can help both parties to identify the effect of each contract term and requirement on the possible result of the contract (Wijk et al., 2011; Settanni et al., 2017).

Chapter 6:   Summary, Contributions, and Suggestions for Future Work

In the process of contract negotiation and contract execution the objectives and constraints of the customer and contract (e.g., the public sector and the private sector) are different. The private sector has full authority to decide how to obtain a maximized long-term profit, specifically, how to control the cost of maintenance while reaching a good performance level so as to receive a better payment. Hence, the profit not only depends on the detailed contract terms, but also on the private sector's Operation and Maintenance (O&M) strategy during the contracted life time of the project. The public sector, on the other hand, is trying to incentivize the private party to sustain a good performance level throughout the contract, given the long-term budget constraint.

This dissertation provides a comprehensive review of the elements of the contract that can help the customer to incentivize the contractor without adding complex terms to the contract requirements. Analytical methods to design maintenance contracts that address the reliability of systems and supply chain operation are reviewed and the existing gaps analyzed. The concept of "Contract Engineering" is introduced as the concurrent design of contract and systems(Kashani-Pour et al., 2017).

A simulation-based method that aids in contract design and negotiation was developed and demonstrated. The method allows the identification of different features of an availability-based contract for a variety of systems that are transitioning from conventional labor and material contract to performance-based contracting.

The simulation method introduced in this dissertation uses different models for the contractor and customer separately which are related through a payment model. The model

aims to guide better design and negotiation of availability contracts and is expected to help both parties understand the essential purpose of the partnership and seek their mutual interest more efficiently. We can extend the model developed in this dissertation (i.e., the controller design) either to target a certain function of availability directly or an availability growth curve through time determined by enforcing more detailed requirements. In our analysis, we assumed that reliability (and thereby demand) is not a control parameter; however, in reality, PBL is designed to incentivize OEMs to improve their reliability (Guajardo et al., 2012) and the effect could be captured in the model. An extension of the model could be used, under certain assumptions, to determine the optimum contract length for such contracts similar to Deng et al. (2015).

The new model has several direct managerial applications: using the model one can quantify that the contractor's prefer a larger assessment interval because: 1) more information is more helpful for demand forecasting, 2) the effect of a few low-performance periods on the overall performance will be minor (more tolerance towards demand variability), and 3) there is more time and opportunity to compensate for a sudden change in demand. The contractor's preferences can be observed by assessing the variance of performance under various assessment intervals (as shown in Figure 30). However, as the simulation model focuses on the expected cost, a long interval will cause more oscillation in contractor behavior and not necessarily help the contractor. Figure 26 shows that by increasing the length of the assessment interval the variability in the performance of the contractor will increase; this variability is measured at the end of each assessment interval. This variability shows the amount of risk the contractor is facing during each period. Moreover, in this work, the contractor does not change their optimal policy after each observation, but one could consider such changes for some contractors. It is also clear that there is no need to increase N_a excessively, and there is a threshold beyond which choosing a larger assessment window size can result in any desired variability for a given performance level.

6.1   Contributions

In general, there has been very little work that links engineering design to the contract design (Sandborn et al., 2017). This dissertation represents the first attempt to formally (and quantitatively) connect these two.

- The approach developed in this dissertation represents a new method for cost modeling and pricing sustainment contracts. To the best of the author's knowledge, no previous cost modeling methods have considered the effect of uncertainty of contractor incentives to the customer cost models and contract requirements selection. This method can be classified as bottom-up discrete-event simulation for cost modeling for outcome-based pricing.

- A comprehensive end-to-end event-based model for modeling the operation, supply-chain and the payment mechanism has been developed using a variety of methods (discrete-event simulation and dynamic programing). This model covers the whole spectrum from the physical layer of the system (represented in the stochastic Petri nets) to the payment structure (based in public-private partnerships) including incentives of maintenance agents and the contractor goal. Such an integrated approach to modeling has not appeared previously in the maintenance and service literature.

- Dynamic Model of Contractor Behavior

Instead of a direct discrete-event simulation, a closed-loop optimal affine-controller has been used to model the predictive and corrective actions of the contractor throughout the inspection periods and total length of contract with response to availability requirements. The affine-controller is used to model an ideal contractor that minimizes the expected total cost of operating the maintenance inventory. Using this assumption the customer can assume a range of uncertainty for the contractor performance and incentives and simulate

the total-life-cost of the contract to find the best set of requirements, i.e., time-assessment window. Usage of such approach in contract modeling is new.

- Payment Plan, Award and Penalty Design

The adaption and extension of "availability payment" concepts currently in use for civil infrastructure PPPs to contract design and pricing for PBL contracts. The model development in this dissertation explores and demonstrates the merit of the civil infrastructure PPP approach for platform-level PBL and new acquisition subsystem PBL contracts. We have focused on availability as the key required outcome and introduce a stochastic and layered availability requirement into the proposed civil infrastructure PPP based PBL contract structure.

- By assuming the cost for the assessments, there will be a trade-off that will provide a global optimum for the number of assessments (N_a) and consequently for the assessment interval. This optimum value of N_a minimizes the cost of operation for the customer. An assessment window that is larger or smaller than this optimum will not benefit the contractor or the customer. However, the cost versus assessment window length relation is not symmetrical around the optimum point and adding more time to the assessment interval has a less detrimental effect than reducing the interval (i.e., assessing more often). In addition, this interval will reduce the cost-risk for the customer as well as the variation of operational cost for the contractor.

- An optimal performance assessment interval was shown to exist that has a considerable influence on the cost of the contract. The concept of an optimal assessment interval has not existed previously in the maintenance service contract literature, but has practical applications for defining contract options periods. This suggests that there is a potential for defining new metrics for contractors that can make availability-based contracts more successful. Moreover, there could be new contract mechanisms designed to account for the time assessment window.

- A multi-objective decision making can benefit from the mean-var analysis to assess contractors with different risk attitudes (risk-averse, risk neutral) with regard to cost-risk and variation of cost at each assessment time.

- The methodology developed in this dissertation can help when a Performance Based Logistic (PBL) Contract is being negotiated. The decision making team can use the data provided in Mean-Var or Cost-Risk analysis and might not need to repeat the simulation during the negotiation given the inclusiveness of the result state-space provided by simulation.

6.2 Future Work

Since the structure is based on simulation-based search, nearly any type of payment model or requirement (level or ramp) can be tested using the developed method. Different payment mechanisms to address the cost variations can be pursued with advanced optimization methods.

Future work can be done using a game theoretic, two-layer optimization structure that models the interaction between the contractor and the customer under sudden changes or interruption in the outcome, or to investigate the effect of sharing of information (PHM information) on total cost. Including system design refreshment decisions as well as requirement trends in time.

The optimization of the performance must be achieved under conditions that include the different types of risk-taking attitudes of the contractor. The role of heuristic search is essential due to complexities in the objective function and operation requirements. For example, objective functions such as conditional value at risk (CVAR) put more focus on the preparedness. The role of uncertainty in demand that is rooted in the reliability or the systems requirements can be further investigated. The physical model that generates such demand can become more complex.

Finally, the Government Accountability Office (GAO, 2016) estimates the current

US weapon portfolio total acquisition cost, including buying and sustaining, will be around $1.44 trillion—and the estimates are based on deliberately optimistic projections. The costs will likely grow, as has happened many times in the past. A data-driven approach can use publicly available sustainment contracts solely based on their base and option period lengths and assess the role of program assessment periods on cost-over runs and availability.

Publications Associated with this Dissertation

Book Chapter

Kashani-Pour, A., Goudarzi, N., Lei, X., & Sandborn, P. (2017). Product-service systems under availability-based contracts: Maintenance optimization and concurrent system and contract design. In Advances in Through-life Engineering Services (pp. 309–329). Springer.

Journal Papers

Jazouli, T., Sandborn, P., & Kashani-Pour, A. (2014). A direct method for determining design and support parameters to meet an availability requirement. International Journal of Performability Engineering, 10(2), 211–225.

Kashani-Pour, A. R., Sandborn, P., & Cui, Q. (2016). Review of quantitative methods for designing availability-based contracts. Journal of Cost Analysis and Parametrics, 9(1), 69–91.

Papers Under Review or to be Submitted

Kashani-Pour, A. & Sandborn, P. (n.d.-a). Optimum requirements selection for contract design in availability-based sustainment acquisitions. To be submitted to Naval Research Logistics.

Kashani-Pour, A. & Sandborn, P. (n.d.-b). Trade-off and sensitivity analysis for new product warranty design using MINLP. To be submitted to Computers and Simulation.

Zhao, X., Kashani-Pour, A. R., & Sandborn, P. (2015). Dynamic performance model for a closed-loop supply chain system with uncertain demand and returns. College Park, MD.

Zhu, X., Kashani-Pour, A., Cui, Q., & Sandborn, P. (n.d.). Contract design for availability payment public-private partnership. To be submitted to International Journal of Project Management.

Peer Reviewed Conference Papers

Bakhshi, R., Sandborn, P., Lei, X., & Kashani-Pour, A. (2015). Return on investment modeling to support cost avoidance business cases for wind farm o&m. In Proceedings of the EWEA Offshore. Copenhagen, Denmark.

Kashani-Pour, A., Zhu, X., Sandborn, P., & Cui, Q. (2014). A new availability-payment model for pricing performance-based logistics contracts. In Proceedings Annual Acquisition Research Symposium.

Lei, X., Goudarzi, N., Kashani-Pour, A., & Sandborn, P. (2016). Incorporation of outcome-based contract requirements in a real options approach for maintenance planning. University of Maryland, College Park College Park United States.

Lei, X., Sandborn, P., Bakhshi, R., & Kashani-Pour, A. (2015a). Development of a maintenance option model to optimize offshore wind farm o&m. In Proceedings of EWEA Offshore. Copenhagen, Denmark.

Lei, X., Sandborn, P., Bakhshi, R., & Kashani-Pour, A. (2015b). Phm based predictive maintenance optimization for offshore wind farms. In 2015 IEEE Conference on Prognostics and Health Management (PHM) (pp. 1–8). IEEE.

Lei, X., Sandborn, P., Goudarzi, N., Bakhashi, R., Kashani-Pour, A., et al. (2015c). Using maintenance options to optimize wind farm o&m. In North American Wind Energy Academy 2015 Symposium. Virginia Tech.

Sandborn, P., Haddad, G., & Kashani-Pour, A. (2014). Development of a maintenance option model to optimize offshore wind farm sustainment. In 32nd ASME Wind Energy Symposium (p. 0362).

Sandborn, P., Kashani-Pour, A., & Goudarzi, N. (2016). Outcome-based contracts - optimizing maintenance contract design for product-service systems. In Proceedings of 5th International Through-Life Engineering Services Conference. Cranfield University.

Appendix A: Affine Controller Design

Consider a discrete-time linear time-varying system, which satisfies the following system transition:

$$x(t + 1) = A(t)x(t) + B(t)u(t) + w(t), \ t = 0, \dots, T - 1 \tag{56}$$

Equation (56) can be rewritten as:

$$x = Gw + Hu + x_0 \tag{57}$$

where

$$G = \begin{bmatrix} 0 & & & & \\ A_1^1 & 0 & & & \\ A_1^2 & A_2^2 & 0 & & \\ & & & \ddots & \\ & & & & 0 \\ A_1^T & A_2^T & \cdots & & A_T^T \end{bmatrix}$$

$$H = \begin{bmatrix} 0 & & & & \\ A_1^1 B(0) & 0 & & & \\ A_1^2 B(0) & A_2^2 B(1) & 0 & & \\ & & & \ddots & \\ A_1^T B(0) & A_2^T B(1) & \cdots & & A_T^T B(T - 1) \end{bmatrix}$$

and

$$x_0 = (x(0), A_0^1 x(0), \dots, A_0^T x(0))^T$$

where $A_\tau^t = A(t-1)A(t-2) \cdots A(\tau)$, and $A_t^t = l$.

Then we consider a casual feedback affine controller, which has the form of:

$$u(t) = \varphi_t(x(0), \dots, x(t)) = u_0(t) + \sum_{\tau=0}^{t} F(t, \tau)x(\tau) \tag{58}$$

φ_t is called the control policy. With a close-loop system, the feedback matrix can be defined as:

$$F = \begin{bmatrix} F(0,0) & & \\ F(1,0) & F(1,1) & \\ & & \\ F(T-1,0) & F(T-1,1) & F(T-1,T-1) \end{bmatrix}$$

Then we will have

$$u = Fx + u_0 \tag{59}$$

With (57) and (59), we can solve for x and u in ters of w, to get

$$\begin{bmatrix} x \\ u \end{bmatrix} = Pw + \begin{pmatrix} \tilde{x} \\ \tilde{u} \end{pmatrix} \tag{60}$$

Where P is called close-loop matrix

$$P = \begin{bmatrix} P_{xw} \\ P_{uw} \end{bmatrix} \tag{61}$$

$$P_{xw} = G + HF(I - HF)^{-1}G$$

$$P_{uw} = F(1 - HF)^{-1}G$$

113

And

$$\tilde{x} = x_0 + Hu_0 + HF(I - HF)^{-1}(x_0 + Hu_0)$$

$$\tilde{u} = F(I - HF)^{-1}(x_0 + Hu_0) + u_0$$

The optimization problem is in general not convex in the design variables F and u_0. By a suitable Q-design procedure, however, these problems can be cast as convex optimization problems, and therefore solved efficiently:

Define:

$$Q = F(I - HF)^{-1} \tag{62}$$

Then

$$F = (I + QH)^{-1}$$

Define

$$r = (I + QH)u_0 \tag{63}$$

Then

$$u = (I + FH)r$$

Then the close-loop matrix P becomes

$$P = \begin{bmatrix} P_{xw} \\ P_{uw} \end{bmatrix} = \begin{bmatrix} (I + HQ)G \\ QG \end{bmatrix} \tag{64}$$

$$\tilde{x} = (I + HQ)x_0 + Hr$$

$$\tilde{u} = Qx_0 + r$$

Therefore:

$$x = (I + HQ)GW + (I + HQ)x_0 + Hr \tag{65}$$

$$u = Q(GW + x_0) + r \tag{66}$$

Appendix B: Literature Review Methods

The strategy used to analyze the existing relevant work on availability contract design and analysis was based on multiple cross-checking of models and contexts of applications and literature. A variety of related literature was studied to identify contract models or contract-oriented applications in the context of performance- and availability-based contacts. Related works were identified through an electronic search of databases that included: Emerald, Science Direct, IEEE Explore, library files and reference lists in relevant papers. In addition, the literature search was extended to the US governement websites (NASA, GOA, DoD, Defense Acquistion University) and consulting companies' web pages (Booz Allen, CSSI, RAND, etc.). In this paper, we only consider works that were foundational, quantitative and explicitly related to designing availability-based contracts and sustainment models.

A structured approach was adopted to develop a framework for the assessment of works from variety of literature from reliability to supply-chain management. This structured approach is based on the following questions: 1) What is the domain of application or theory the paper is focused on? 2) What is the context of the problem statement in the paper? 3) What type of contract is being addressed? 4) If the paper is in the reliability domain, what elements relevant to contract design are discussed in the paper? 5) If the paper is in the supply chain realm, what is the modeling contribution of the paper? 6) Is the work from the view point of the customer or the contactor (or both)? 7) What is the quantitative nature of the model? And 8) If an optimization method is used, how are the constraints or objectives connected to contract requirements? In addition, there was a

careful study of the quantitative problem statement in each optimization paper including objective function, constraints, time horizon of the algorithm (finite, infinite) and the scope of the model (fleet level or individual unit of the system). This approach is shown with the data collected in this study in Table B.1 for a selected group of papers. The organization of the paper is based on the framework shown in Table B.1.

Table B.1: Framework for comparison of examples of computational approaches to availability-based contract design

	ELEMENTS OF CONTRACT			CONTEXT AND APPLICATION OF CONTRACT			METHODS AND MODELS		
Type of paper	**Type of Contract**	**View Point**	**Requirements Constraints**	**Context**	**RAMS Aspect**	**Supply Chain Aspect**	**Modeling Approach**	**Time Horizon**	**Optimization Method**
Review Report Case-study Method Application	PBL/ Availability-based Contract/ PPP	Contractor Customer Both	Objective function Deterministic/ Probabilistic Fleet Level/ Product-unit	Economics OR L+SCM RAMS LCC PM ACQ	Reliability Inventory Availability Maintenance CM Warranty LCC	Logistics Inventory Transportation Supply Chain	Markov Models/ Petri-nets/ DES /Close-Form SD/ FL/ Simulation	Periodical Infinite (steady-state) Finite	LP DP Stochastic Programing GA Heuristics Monte-Carlo
Datta and Roy, 2010	Review of 87 papers	Availability-based Contract	Both		LCC	LCC			
Jazouli et al., 2010	Method	Availability-based Contract	Contractor	Deterministic / Product unit	RAMS LCC	LCC		DES	Finite Parametric	Monte-Carlo
Dandotiya, et al., 2008	Method	Availability-based Contract		Deterministic	RAMS	Maintenance		Closed-form	Finite parametric	LP
Ferguson et al., 2011	Method	PBL	Both		SCM ACQ		Inventory	Closed Form		Monte-Carlo
de la Garza, et al., 2008	Method	PBL	Customer	Deterministic		CM	road condition		Periodical Parametric	Sampling Procedure
Mendoza et al., 2005	Report	PBL	Customer		PM					
Komoto, et al., 2011	Method	PBL	Contractor		LCC			DES		Stochastic Programing
Volvoi, et al., 2004	Method			Fleet Level		Reliability		Petri Nets		
Arora et al., 2010	Method		Contractor		OR		Inventory supply chain		Finite Parametric	Heuristic
Bowman & Schmee, 2001	Method		Contractor		OR	LCC Availability		DES	5-20 years	DES
Nilsson et al., 2007	Method Case-Study		Contractor		LCC	LCC		Closed-form	20 years	
Frangopol et al., 2003	Review of 195 papers		Both		LCC RAMS	Warranty RAMS	Inventory			
Lapa, et al., 2006	Method			Cost& Reliability		PM		Probabilistic Closed-form	Flexible intervals	GA

ACQ : Acquisition ; CM: Condition Monitoring; DES: Discrete-Event Simulation DP: Dynamic Programing ; FL: Fuzzy Logic; GA: Genetic Algorithm; LCC: Life Cycle Cost; LP : Linear Programing; L+SCM : Logistics and Supply Chain Management; OR Operational Research ; PBL: Performance Based Contract; PM: Predicative Maintenance; PPP: Public Private Partnership; RAMS: Reliability, Availability, Maintainability Safety; SD: System Dynamics.

References

Aalst, W. M. van der, Lohmann, N., Massuthe, P., Stahl, C., & Wolf, K. (2010). Multiparty contracts: Agreeing and implementing interorganizational processes. The Computer Journal, 53(1), 90–106.

Albright, S. C. & Soni, A. (1988). Markovian multiechelon repairable inventory system. Naval Research Logistics (NRL), 35(1), 49–61.

Alfredsson, P. (1997). Optimization of multi-echelon repairable item inventory systems with simultaneous location of repair facilities. European journal of operational research, 99(3), 584–595.

Alrabghi, A. & Tiwari, A. (2015). State of the art in simulation-based optimisation for maintenance systems. Computers & Industrial Engineering, 82, 167–182.

Angerhofer, B. J. & Angelides, M. C. (2000). System dynamics modelling in supply chain management: Research review. In Simulation conference, 2000. proceedings. winter (Vol. 1, pp. 342–351). IEEE.

Arora, V., Chan, F. T., & Tiwari, M. K. (2010). An integrated approach for logistic and vendor managed inventory in supply chain. Expert Systems with Applications, 37(1), 39–44.

Ashgarizadeh, E. & Murthy, D. (2000). Service contracts: A stochastic model. Mathematical and Computer Modelling, 31(10-12), 11–20.

Austin, M., Mayank, V., & Shmunis, N. (2006). Paladinrm: Graph-based visualization of requirements organized for team-based design. Systems Engineering, 9(2), 129–145.

Azarm, S. & Mourelatos, Z. P. (2006). Robust and reliability-based design. Journal of Mechanical Design, 128(4), 829–831.

Baines, T., Lightfoot, H. W., & Kay, J. M. (2009). Servitized manufacture: Practical challenges of delivering integrated products and services. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 223(9), 1207–1215.

Bakhshi, R. & Sandborn, P. (2017). A return on investment model for the implementation of new technologies on wind turbines. IEEE Transactions on Sustainable Energy, PP(99), 1–1.

Bakhshi, R., Sandborn, P., Lei, X., & Kashani-Pour, A. (2015). Return on investment modeling to support cost avoidance business cases for wind farm o&m. Proceedings of the EWEA Offshore.

Bangemann, T., Rebeuf, X., Reboul, D., Schulze, A., Szymanski, J., Thomesse, J.-P., … Zerhouni, N. (2006). Proteus—creating distributed maintenance systems through an integration platform. Computers in Industry, 57(6), 539–551.

Barata, J., Soares, C. G., Marseguerra, M., & Zio, E. (2002). Simulation modelling of repairable multi-component deteriorating systems for 'on condition' maintenance optimisation. Reliability Engineering & System Safety, 76(3), 255–264.

Barlow, R. E. & Proschan, F. (1996). Mathematical theory of reliability. p. 280. SIAM.

Benvenuti, L., Ferrari, A., Mangeruca, L., Mazzi, E., Passerone, R., & Sofronis, C. (2008a). A contract-based formalism for the specification of heterogeneous systems. In Specification, Verification and Design Languages, 2008. FDL 2008. Forum on (pp. 142–147). IEEE.

Benvenuti, L., Ferrari, A., Mazzi, E., & Vincentelli, A. S. (2008b). Contract-based design for computation and verification of a closed-loop hybrid system. In International Workshop on Hybrid Systems: Computation and Control (pp. 58–71). Springer.

Block, J., Ahmadi, A., & Tyrberg, T. (2014). Using monte carlo simulation as support for decision making while negotiating a pbl contract. In Aerospace Conference, 2014 IEEE (pp. 1–7). IEEE.

Block, J., Tyrberg, T., & Fuquing, Y. (2010). Optimal repair for repairable components during phaseout an aircraft fleet. In Aerospace Conference, 2010 IEEE (pp. 1–8). IEEE.

Bodner, D., Smith, R., & Rouse, B. (2009). Simulation-based decision support for acquisition policy and process design: The effect of system and enterprise characteristics on acquisition outcomes. DTIC Document.

Borgonovo, E., Marseguerra, M., & Zio, E. (2000). A monte carlo methodological approach to plant availability modeling with maintenance, aging and obsolescence. Reliability Engineering & System Safety, 67(1), 61–73.

Bowman, R. A. & Schmee, J. (2001). Pricing and managing a maintenance contract for a fleet of aircraft engines. Simulation, 76(2), 69–77.

Boyce, J. & Banghart, A. (2012). Performance based logistics and project proof point. Defense AT&L: Product Support Issue, 41(2), 26–30.

Brailsford, S. C., Desai, S. M., & Viana, J. (2010). Towards the holy grail: Combining system dynamics and discrete-event simulation in healthcare. In Simulation Conference (WSC), Proceedings of the 2010 Winter (pp. 2293–2303). IEEE.

Brown, M. & Hockly, C. (2001). Cost of specifying maintenance/failure free operating periods for royal air force aircraft. In Reliability and Maintainability Symposium, 2001. Proceedings. Annual (pp. 425–432). IEEE.

Cachon, G. P. (2003). Supply chain coordination with contracts. Handbooks in operations research and management science, 11, 227–339.

Cachon, G. P. & Lariviere, M. A. (2005). Supply chain coordination with revenue-sharing contracts: Strengths and limitations. Management science, 51(1), 30–44.

Cai, X. & Tyagi, S. (2014). Development of a product life-cycle cost estimation model to support engineering decision-making in a multi-generational product development environment. Journal of Cost Analysis and Parametrics, 7(3), 219–235.

Canfield, R. (1986). Cost optimization of periodic preventive maintenance. IEEE Transactions on Reliability, 35(1), 78–81.

Caro, J. J. (2005). Pharmacoeconomic analyses using discrete event simulation. Pharmacoeconomics, 23(4), 323–332.

Caro, J. J., Möller, J., & Getsios, D. (2010). Discrete event simulation: The preferred technique for health economic evaluations? Value in health, 13(8), 1056–1060.

Carroll, J., McDonald, A., Barrera Martin, O., McMillan, D., & Bakhshi, R. (2015). Offshore wind turbine sub-assembly failure rates through time. In Scientific Proceedings (pp. 112–116). European Wind Energy Association.

Carvalho, H., Barroso, A. P., Machado, V. H., Azevedo, S., & Cruz-Machado, V. (2012). Supply chain redesign for resilience using simulation. Computers & Industrial Engineering, 62(1), 329–341.

Chen, H., Labadi, K., & Amodeo, L. (2006). Modeling, analysis, and optimization of logistics systems petri net based approaches. In Service Systems and Service Management, 2006 International Conference on (Vol. 1, pp. 575–582). IEEE.

Chen, X., Li, L., & Zhou, M. (2012). Manufacturer's pricing strategy for supply chain with warranty period-dependent demand. Omega, 40(6), 807–816.

Cushway, G. (2006). Availability contracting. Defence Management Journal, 32(1), 74–75.

Cuthbertson, R. & Piotrowicz, W. (2011). Performance measurement systems in supply chains: A framework for contextual analysis. International Journal of Productivity and Performance Management, 60(6), 583–602.

Damm, W., Josko, B., & Peikenkamp, T. (2009). Contract based ISO CD 26262 safety analysis. Warrendale, PA. available at: http://papers.sae.org/2009-01-0754/ (accessed 22 March 2013).

Dandotiya, R., Fuqing, Y., & Kumar, U. (2008). Optimal maintenance decision for line reparable units (lru) for an aircraft system: A conceptual approach. Quarterly Journal of the Operational Research Society of India (OPSEARCH), 45(3), 291–302.

Datta, P. P. & Roy, R. (2010). Cost modelling techniques for availability type service support contracts: A literature review and empirical study. CIRP Journal of Manufacturing Science and Technology, 3(2), 142–157.

Defense Acquisition University Press. (2005). Performance based logistics: A program manager's product support guide. Ft. Belvoir, Va: Defense Acquisition University Press.

Dekker, R. (1996). Applications of maintenance optimization models: A review and analysis. Reliability Engineering & System Safety, 51(3), 229–240.

Dekker, R. & Scarf, P. A. (1998). On the impact of optimisation models in maintenance decision making: The state of the art. Reliability Engineering & System Safety, 60(2), 111–119.

Deng, Q., Zhang, L., Cui, Q., & Jiang, X. (2015). A simulation-based decision model for designing contract period in building energy performance contracting. Building and Environment, 71, 71–80.

Dijkhuizen, G. van & Van der Heijden, M. (1999). Preventive maintenance and the interval availability distribution of an unreliable production system. Reliability Engineering & System Safety, 66(1), 13–27.

Disney, S. M. & Towill, D. R. (2002). A procedure for the optimization of the dynamic response of a vendor managed inventory system. Computers & industrial engineering, 43(1), 27–58.

Doerr, K., Lewis, I., & Eaton, D. R. (2005). Measurement issues in performance-based logistics. Journal of Public Procurement, 5(2), 164.

Dohi, T., Kaio, N., Young Yun, W., Bertolini, M., Bevilacqua, M., & Mason, G. (2006). Reliability design of industrial plants using petri nets. Journal of Quality in Maintenance Engineering, 12(4), 397–411.

Dulcos, R. & Shepherd, N. (1991). Repair level analysis software (LSA Subtask 303.2. 7), User's Manual (Version 1.0). American Power Jet Co Ridgefield NJ.

Eijkenaar, F., Emmert, M., Scheppach, M., & Schöffski, O. (2013). Effects of pay for performance in health care: A systematic review of systematic reviews. Health policy, 110(2), 115–130.

Erkoyuncu, J. A. (2011). Cost uncertainty management and modelling for industrial product-service systems (Doctoral dissertation, Cranfield University).

Erkoyuncu, J. A., Roy, R., Shehab, E., & Wardle, P. (2009). Uncertainty challenges in service cost estimation for product-service systems in the aerospace and defence industries. In Proceedings of the 19th CIRP Design Conference–Competitive Design. Cranfield University Press.

Federal Acquisition Regulation (FAR) § 16.505(c)(1), a limitation on task order contracts for advisory and assistance services. (2010).

Federal Acquisition Regulation (FAR) § 17.104(a), a limitation on multi-year contracts. (2010).

Federal Acquisition Regulation (FAR) § 17.204(e), a limitation on contracts with options. (2010).

Federal Acquisition Regulation (FAR) § 22.1002-1, a limitation on contracts covered by the Service Contract Act of 1965, as amended, 41 U.S.C. § 353(d). (2010).

Federgruen, A. & Zheng, Y.-S. (1992). An efficient algorithm for computing an optimal (r, q) policy in continuous review stochastic inventory systems. Operations research, 40(4), 808–813.

Feeney, G. J. & Sherbrooke, C. C. (1966). The (s- 1, s) inventory policy under compound poisson demand. Management Science, 12(5), 391–411.

Feldman, K., Jazouli, T., & Sandborn, P. A. (2009). A methodology for determining the return on investment associated with prognostics and health management. IEEE Transactions on Reliability, 58(2), 305–316.

Ferguson, J. & Sodhi, M. (2011). Using simulation for setting terms of performance based contracts. In Operations Research Proceedings 2010 (pp. 465–470). Springer.

Ferreira, R. J., Almeida, A. T. de, & Cavalcante, C. A. (2009). A multi-criteria decision model to determine inspection intervals of condition monitoring based on delay time analysis. Reliability Engineering & System Safety, 94(5), 905–912.

Fincher, J. H. (2016). What are we waiting for? customer wait time, fill-rate, and marine corps equipment operational availability (Doctoral dissertation, Monterey, California: Naval Postgraduate School).

Frangopol, D. M. & Maute, K. (2003). Life-cycle reliability-based optimization of civil and aerospace structures. Computers & structures, 81(7), 397–410.

French, E. & Jones, J. B. (2004). On the distribution and dynamics of health care costs. Journal of Applied Econometrics, 19(6), 705–721.

Fu, M. C. (1994). Optimization via simulation: A review. Annals of operations research, 53(1), 199–247.

Gansler, J. S. (2000). Memorandum on performance-based services acquisition (pbsa).

Gansler, J. S. & Lucyshyn, W. (2006). Evaluation of performance based logistics. Naval Postgraduate School.

GAO. (2002). Guidance needed for using performance-based service contracting. *GAO-02-1049*: Published: Sep 23, 2002. Publicly Released: Oct 23, 2002.

GAO. (2005). Dod needs to demonstrate that performance-based logistics contracts are achieving expected benefits. *GAO-05-966*: Published: Sep 9, 2005. Publicly Released: Sep 9, 2005.

GAO. (2016). Defense acquisitions: Assessments of selected weapon programs. *GAO-16-329SP*: Published: March 2016.

Garg, S., Puliafito, A., Telek, M., & Trivedi, K. S. (1995). Analysis of software rejuvenation using markov regenerative stochastic petri net. In Software Reliability Engineering, 1995. Proceedings., Sixth International Symposium on (pp. 180–187). IEEE.

Garza, J. de la, Pinero, J., & Ozbek, M. (2006). A framework for monitoring performance-based road maintenance contracts.

Garza, J. de la, Piñero, J., & Ozbek, M. (2008). Sampling procedure for performance-based road maintenance evaluations. Transportation Research Record: Journal of the Transportation Research Board, (2044), 11–18.

Glas, A. H., Hofmann, E., & Michael, E. (2013). Performance-based logistics: A portfolio for contracting military supply. International Journal of Physical Distribution & Logistics Management, 43(2), 2.

Godoy, O., Kraft, N. J., & Levine, J. M. (2014). Phylogenetic relatedness and the determinants of competitive outcomes. Ecology Letters, 17(7), 836–844.

Goebel, K., Krok, M., & Sutherland, H. (2000). Diagnostic information fusion: Requirements flowdown and interface issues. In Aerospace Conference Proceedings, 2000 IEEE (Vol. 6, pp. 155–162). IEEE.

Grossmann, I. E. (2012). Advances in mathematical programming models for enterprise-wide optimization. Computers & Chemical Engineering, 47, 2–18.

Gruneberg, S., Hughes, W., & Ancell, D. (2007). Risk under performance-based contracting in the uk construction sector. Construction Management and Economics, 25(7), 691–699.

Guajardo, J. A., Cohen, M. A., Kim, S.-H., & Netessine, S. (2012). Impact of performance-based contracting on product reliability: An empirical analysis. Management Science, 58(5), 961–979.

Gupta, D., Vedantam, A., & Azadivar, J. (2011). Optimal contract mechanism design for performance-based contracts. Minnesota Department of Transportation Research Services Section.

Gupta, O. K. & Ravindran, A. (1985). Branch and bound experiments in convex nonlinear integer programming. Management science, 31(12), 1533–1546.

Haas, P. J. & Shedler, G. S. (1986). Regenerative stochastic petri nets. Performance Evaluation, 6(3), 189–204.

Hamidi, M., Liao, H., & Szidarovszky, F. (2014). A game-theoretic model for outsourcing maintenance services. In Reliability and Maintainability Symposium (RAMS), 2014 Annual (pp. 1–6). IEEE.

Hansen, S. J. (2006). Performance contracting: Expanding horizons. The Fairmont Press, Inc.

Hawker, G. S. & McMillan, D. A. (2015). The impact of maintenance contract arrangements on the yield of offshore wind power plants. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 229(5), 394–402.

Hershey, P. C. (1996). High availability data processing system and method using finite state machines. US Patent 5,544,077. Google Patents.

Hockley, C. J. & Appleton, D. (1997). Setting the requirements for the royal air force's next generation aircraft. In Reliability and Maintainability Symposium. 1997 Proceedings, Annual (pp. 44–49). IEEE.

Hockley, C. J., Smith, J. C., & Lacey, L. J. (2011). Contracting for availability and capability in the defence environment. In Complex engineering service systems (pp. 237–256). Springer.

Hokstad, P., Langseth, H., Lindqvist, B. H., & Vatn, J. (2005). Failure modeling and maintenance optimization for a railway line. International Journal of Performability Engineering, 1(1), 51.

Hong, S., Wernz, C., & Stillinger, J. D. (2016). Optimizing maintenance service contracts through mechanism design theory. Applied Mathematical Modelling, 40(21), 8849–8861.

Hooper, L. (2008). Paying for performance: Uncertainty, asymmetric information and the payment model. Research in Transportation Economics, 22(1), 157–163.

Hussain, M., Drake, P. R., & Myung Lee, D. (2012). Quantifying the impact of a supply chain's design parameters on the bullwhip effect using simulation and taguchi design of experiments. International Journal of Physical Distribution & Logistics Management, 42(10), 947–968.

Immonen, A. & Niemelä, E. (2008). Survey of reliability and availability prediction methods from the viewpoint of software architecture. Software and Systems Modeling, 7(1), 49–65.

Janakiraman, G., Santos, J. R., & Turner, Y. (2004). Automated system design for availability. In Dependable Systems and Networks, 2004 International Conference on (pp. 411–420). IEEE.

Jazouli, T., Sandborn, P., & Kashani-Pour, A. (2014). A direct method for determining design and support parameters to meet an availability requirement–parameters affecting both downtime and uptime. International Journal of Performability Engineering, 10(6), 649.

Jazouli, T. & Sandborn, P. (2010). A design for availability approach for use with phm. In Annual Conference of the Prognostics and Health Management Society (Vol. 2010, pp. 1–12).

Jazouli, T. & Sandborn, P. (2011). Using phm to meet availability-based contracting requirements. In Prognostics and Health Management (PHM), 2011 IEEE Conference on (pp. 1–12). IEEE.

Jian, L. & Hong-Fu, Z. (2004). The predictive models of maintenance costs for a civil airplane. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 218(5), 347–351.

Jin, M. & Wu, S. D. (2002). Procurement auctions with supplier coalitions: Validity requirements and mechanism design. Relation, 10(1.70), 2759–2883.

Jin, T., Tian, Z., & Xie, M. (2015). A game-theoretical approach for optimizing maintenance, spares and service capacity in performance contracting. International Journal of Production Economics, 161, 31–43.

Jin, T. & Wang, P. (2011). Planning performance based logistics considering reliability and usage uncertainty. Journal of the Operational Research Society, 10(63), 1467–1478.

Johnstone, S., Dainty, A., & Wilkinson, A. (2008). In search of 'product-service': Evidence from aerospace, construction, and engineering. The Service Industries Journal, 28(6), 861–875.

Juang, Y.-S., Lin, S.-S., & Kao, H.-P. (2008). A knowledge management system for series-parallel availability optimization and design. Expert Systems with Applications, 34(1), 181–193.

Kajal, S., Tewari, P., & Saini, P. (2013). Availability optimization for coal handling system using genetic algorithm. International Journal of Performability Engineering, 9(1), 109–116.

Karnon, J. (2003). Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. Health economics, 12(10), 837–848.

Kashani-Pour, A. R., Sandborn, P., & Cui, Q. (2016). Review of quantitative methods for designing availability-based contracts. Journal of Cost Analysis and Parametrics, 9(1), 69–91.

Kashani-Pour, A., Goudarzi, N., Lei, X., & Sandborn, P. (2017). Product-service systems under availability-based contracts: Maintenance optimization and concurrent system and contract design. In L. Redding, R. Roy, & A. Shaw (Eds.), Advances in Through-life Engineering Services (pp. 309–329). Springer.

Kashani-Pour, A., Zhu, X., Sandborn, P., & Cui, Q. (2014). A new availability-payment model for pricing performance-based logistics contracts. In Proceedings Annual Acquisition Research Symposium.

Kastsian, D. & Mönnigmann, M. (2011). Optimization of a vendor managed inventory supply chain with guaranteed stability and robustness. International Journal of Production Economics, 131(2), 727–735.

Keskin, B. B., Melouk, S. H., & Meyer, I. L. (2010). A simulation-optimization approach for integrated sourcing and inventory decisions. Computers & Operations Research, 37(9), 1648–1661.

Kim, B. & Park, S. (2008). Optimal pricing, eol (end of life) warranty, and spare parts manufacturing strategy amid product transition. European Journal of Operational Research, 188(3), 723–745.

Kim, S.-H., Cohen, M. A., & Netessine, S. (2007). Performance contracting in after-sales service supply chains. Management Science, 53(12), 1843–1858.

Kim, W., Yang, J., & Ahn, S. (2009). Determining the periodic maintenance interval for guaranteeing the availability of a system with a linearly increasing hazard rate: Industry application. International Journal of Industrial Engineering: Theory, Applications and Practice, 16(2), 126–134.

Knight, J. T. & Singer, D. J. (2014). Applying real options analysis to naval ship design. Proceedings of ASNE Day.

Kobren, B. (2009). What performance based logistics is and what it is not–and what it can and cannot do. Defense AR Journal, 16(3), 254–268.

Kohani, M. & Pecht, M. (2015). New minimum relative humidity requirements are expected to lead to more medical device failures. Journal of Medical Systems, 40(3), 58.

Komoto, H., Tomiyama, T., Silvester, S., & Brezet, H. (2011). Analyzing supply chain robustness for oems from a life cycle perspective using life cycle simulation. International Journal of Production Economics, 134(2), 447–457.

Labadi, K., Chen, H., & Amodeo, L. (2007). Modeling and performance evaluation of inventory systems using batch deterministic and stochastic petri nets. IEEE Transac-

tions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 37(6), 1287–1302.

Lafontaine, F. & Slade, M. E. (2002). Incentive contracting and the franchise decision. In Game theory and business applications (pp. 133–188). Springer.

Lai, C., Xie, M., Poh, K.-L., Dai, Y.-S., & Yang, P. (2002). A model for availability analysis of distributed software/hardware systems. Information and software technology, 44(6), 343–350.

Lapa, C. M. F., Pereira, C. M. N., & Barros, M. P. de. (2006). A model for preventive maintenance planning by genetic algorithms based in cost and reliability. Reliability Engineering & System Safety, 91(2), 233–240.

Lei, X. & Sandborn, P. A. (2016). Phm-based wind turbine maintenance optimization using real options. Int J Progn Health Manag, 7(1), 1–14.

Lei, X. & Sandborn, P. A. (2017). Maintenance scheduling based on remaining useful life predictions for wind farms managed using power purchase agreements. Renewable Energy.

Levitt, T. (1972). Production-line approach to service. Harvard Business Review, 50(5), 41–52.

Li, M., Azarm, S., & Boyars, A. (2006). A new deterministic approach using sensitivity region measures for multi-objective robust and feasibility robust design optimization. Journal of mechanical design, 128(4), 874–883.

Liao, S.-H., Hsieh, C.-L., & Lai, P.-J. (2011). An evolutionary approach for multi-objective optimization of the integrated location–inventory distribution network problem in vendor-managed inventory. Expert Systems with Applications, 38(6), 6768–6776.

Lindqvist, B. H. (2006). On the statistical modeling and analysis of repairable systems. Statistical science, 532–551.

Markowitz, H. (1952). Portfolio selection. The journal of finance, 7(1), 77–91.

Marseguerra, M. & Zio, E. (2000). Optimizing maintenance and repair policies via a combination of genetic algorithms and monte carlo simulation. Reliability Engineering & System Safety, 68(1), 69–83.

Marseguerra, M., Zio, E., & Podofillini, L. (2005). Multiobjective spare part allocation by means of genetic algorithms and monte carlo simulation. Reliability Engineering & System Safety, 87(3), 325–335.

Materiel Availability. (2010). Available at: https://dap.dau.mil/acquipedia/Pages/ArticleDetails.aspx?aid=462992de-049e-40a4-812c-44713308e1aa (accessed 30 July 2014).

McCall, J. J. (1965). Maintenance policies for stochastically failing equipment: A survey. Management science, 11(5), 493–524.

Meier, H., Roy, R., & Seliger, G. (2010). Industrial product-service systems—ips 2. CIRP Annals-Manufacturing Technology, 59(2), 607–627.

Memarzadeh, M., Pozzi, M., & Zico Kolter, J. (2014). Optimal planning and learning in uncertain environments for the management of wind farms. Journal of Computing in Civil Engineering, 29(5), 04014076.

Mendoza, K. A. & Devlin, L. A. (2005). Performance based logistics and the implications of organizational design. Monterey, California. Naval Postgraduate School.

Mickens, J. W. & Noble, B. D. (2006). Exploiting availability prediction in distributed systems. Ann Arbor, 1001, 48103.

MIL-PRF-5041K. (1998). Performance specification: Tires, ribbed tread, pneumatic, aircraft. available at: http://everyspec.com/MIL-PRF/MIL-PRF-000100-09999/MIL-PRF-5041K_8354/.

MIL-PRF-7726K. (2007). Performance specification: Retread tries, ribbed tread, pneumatic aircraft. available at: http://everyspec.com/MIL-PRF/MIL-PRF-000100-09999/MIL-PRF-7726K_SUPPLEMENT-1_8353/.

Möller, B., Liebscher, M., Pannier, S., Graf, W., & Sickert, J.-U. (2011). An inverse solution of the lifetime-oriented design problem. Structure and Infrastructure Engineering, 7(5), 325–340.

Neuman, C. & Bonhomme, N. (1974). Evaluation of maintenance policies using markov chains and fault tree analysis. IEEE Transactions on reliability, 24(1), 37–44.

Newsome, J. (2008). Study reveals performance based logistics contract activity continues to grow, despite industry challenges. Aviation Week Performance-Based Logistics Survey, available at: http://thecenter.utk.edu/images/Users/1/PBL/WhitePapers/AviationWeekPBLSurveyConductedforSAPAmericas.pdf.

Nilsson, J. & Bertling, L. (2007). Maintenance management of wind power systems using condition monitoring systems—life cycle cost analysis for two case studies. IEEE Transactions on energy conversion, 22(1), 223–229.

Nowicki, D. R., Kumar, U., Steudel, H., & Verma, D. (2008). Spares provisioning under performance-based logistics contract: Profit-centric approach. Journal of the Operational Research Society, 59(3), 342–352.

Nowicki, D. R., Randall, W. S., & Ramirez-Marquez, J. E. (2012). Improving the computational efficiency of metric-based spares algorithms. European journal of operational research, 219(2), 324–334.

Nuzzo, P., Xu, H., Ozay, N., Finn, J. B., Sangiovanni-Vincentelli, A. L., Murray, R. M., … Seshia, S. A. (2014). A contract-based methodology for aircraft electric power system design. IEEE Access, 2, 1–25.

Ozbek, M. E. & Jesus, M. (2007). Development of performance warranties for performance-based road maintenance contracts. Alternative Project Delivery Procurement, and Contracting Methods for Highways, 20–36.

Pinero, J. C. (2003). A framework for monitoring performance-based road maintenance (Doctoral dissertation, Virginia Tech).

Plank, J. S. & Thomason, M. G. (1999). The average availability of parallel checkpointing systems and its importance in selecting runtime parameters. In Fault-Tolerant Computing, 1999. Digest of Papers. Twenty-Ninth Annual International Symposium on (pp. 250–257). IEEE.

Randall, W. S., Nowicki, D. R., & Hawkins, T. G. (2011). Explaining the effectiveness of performance-based logistics: A quantitative examination. The International Journal of Logistics Management, 22(3), 324–348.

Ratzer, A. V., Wells, L., Lassen, H. M., Laursen, M., Qvortrup, J. F., Stissing, M. S., … Jensen, K. (2003). Cpn tools for editing, simulating, and analysing coloured petri nets. In International Conference on Application and Theory of Petri Nets (pp. 450–462). Springer.

Ren, Z. J., Cohen, M. A., Ho, T. H., & Terwiesch, C. (2006). Sharing forecast information in a long-term supply chain relationship. Operations research.

Rodrigues, D., Erkoyuncu, J., Starr, A., Wilding, S., Dibble, A., & Laity, M. (2015). Review of the modelling approaches for availability contracts in the military context. Procedia CIRP, 30, 451–456.

Rodriguez, M. A. & Vecchietti, A. (2010). Inventory and delivery optimization under seasonal demand in the supply chain. Computers & Chemical Engineering, 34(10), 1705–1718.

Sandborn, P., Haddad, G., & Kashani-Pour, A. (2014). Development of a maintenance option model to optimize offshore wind farm sustainment. In 32nd ASME Wind Energy Symposium (p. 0362).

Sandborn, P., Kashani-Pour, A., Goudarzi, N., & Lei, X. (2017). Outcome-based contracts– towards concurrently designing products and contracts. Procedia CIRP, 59, 8–13.

Sato, N. & Trivedi, K. S. (2007). Accurate and efficient stochastic reliability analysis of composite services using their compact markov reward model representations. In

Services Computing, 2007. SCC 2007. IEEE International Conference On (pp. 114–121). IEEE.

Scherer, F. M. (1964). The theory of contractual incentives for cost reduction. The Quarterly Journal of Economics, 78(2), 257–280.

Schuëller, G. I. & Jensen, H. A. (2008). Computational methods in optimization considering uncertainties–an overview. Computer Methods in Applied Mechanics and Engineering, 198(1), 2–13.

Settanni, E., Thenent, N. E., Newnes, L. B., Parry, G., & Goh, Y. M. (2017). Mapping a product-service-system delivering defence avionics availability. International Journal of Production Economics, 186, 21–32.

Sharma, D., Cui, Q., Chen, L., & Lindly, J. (2010). Balancing private and public interests in public-private partnership contracts through optimization of equity capital structure. Transportation Research Record: Journal of the Transportation Research Board, (2151), 60–66.

Sherbrooke, C. C. (1971). An evaluator for the number of operationally ready aircraft in a multilevel supply system. Operations Research, 19(3), 618–635.

Sherbrooke, C. C. (2006). Optimal inventory modeling of systems: Multi-echelon techniques. Springer Science & Business Media.

Sherif, Y. & Smith, M. (2006). Optimal maintenance models for systems subject to failure–a review. Naval Research Logistics Quarterly, 28(1), 47–74.

Simpson, K. N., Strassburger, A., Jones, W. J., Dietz, B., & Rajagopalan, R. (2009). Comparison of Markov model and discrete-event simulation techniques for HIV. Pharmacoeconomics, 27(2), 159–165.

Skaf, J. & Boyd, S. P. (2010). Design of affine controllers via convex optimization. IEEE Transactions on Automatic Control, 55(11), 2476–2487.

So, K. C. & Tang, C. S. (2000). Modeling the impact of an outcome-oriented reimbursement policy on clinic, patients, and pharmaceutical firms. Management Science, 46(7), 875–892.

Sols, A., Nowicki, D., & Verma, D. (2008). N-dimensional effectiveness metric-compensating reward scheme in performance-based logistics contracts. Systems Engineering, 11(2), 93–106.

Sun, X., Nuzzo, P., Wu, C.-C., & Sangiovanni-Vincentelli, A. (2009). Contract-based system-level composition of analog circuits. In Design Automation Conference, 2009. DAC'09. 46th ACM/IEEE (pp. 605–610). IEEE.

Szentiványi, D. (2005). Performance studies of fault-tolerant middleware (Doctoral dissertation, Institutionen för datavetenskap).

Szentiványi, D., Nadjm-Tehrani, S., & Noble, J. M. (2005). Optimal choice of checkpointing interval for high availability. In Dependable Computing, 2005. Proceedings. 11th Pacific Rim International Symposium on (8–pp). IEEE.

Taboada, H. A., Espiritu, J. F., & Coit, D. W. (2008). Moms-ga: A multi-objective multi-state genetic algorithm for system reliability optimization design problems. IEEE Transactions on Reliability, 57(1), 182–191.

Thompson, L. B. (2010). The real reason defense companies hate fixed-price r&d contracts. Blog. Available at: http://www.lexingtoninstitute.org/the-real-reason-defense-companies-hate-fixed-price-rd-contracts.

Trefethen, L. N. & Bau III, D. (1997). Numerical linear algebra. Siam.

Tsay, A. A. (1999). The quantity flexibility contract and supplier-customer incentives. Management science, 45(10), 1339–1358.

U.S. Government Accountability Office. (2008). Improved analysis and cost data needed to evaluate the cost-effectiveness of performance-based logistics.

Verma, A. K., Srividya, A., & Gaonkar, R. S. P. (2007). Maintenance and replacement interval optimization using possibilistic approach. International Journal of Modelling and Simulation, 27(2), 193–199.

Volovoi, V. (2004a). Modeling multiphased missions using stochastic petri nets with aging tokens. In Reliability and Maintainability, 2004 Annual Symposium-RAMS (pp. 232–236). IEEE.

Volovoi, V. (2004b). Modeling of system reliability petri nets with aging tokens. Reliability Engineering & System Safety, 84(2), 149–161.

Volovoi, V. (2006). Stochastic petri nets modeling using spn@. In Reliability and Maintainability Symposium, 2006. RAMS'06. Annual (pp. 75–81). IEEE.

Volovoi, V., Kavalieratos, G., Waters, M., & Mavris, D. (2004). Modeling the reliability of distribution systems using petri nets. In Harmonics and Quality of Power, 2004. 11th International Conference on (pp. 567–572). IEEE.

Wang, W. (2010). A model for maintenance service contract design, negotiation and optimization. European Journal of Operational Research, 201(1), 239–246.

Wei, C., Li, Y., & Cai, X. (2011). Robust optimal policies of production and inventory with uncertain returns and demand. International Journal of Production Economics, 134(2), 357–367.

Whitehead, T. & Jagdale, S. (2008). Pricing performance based contracts: A risk-based approach.

Wijk, O., Olinger, T., & Hell, R. (2011). Simulation as support for decision making in negotiations. In 14th annual System Engineering Conference in San Diego, USA. Available at: www.systecon.se/documents/Simulation_as_support.pdf.

Wu, C.-C., Chou, C.-Y., & Huang, C. (2007). Optimal burn-in time and warranty length under fully renewing combination free replacement and pro-rata warranty. Reliability Engineering & System Safety, 92(7), 914–920.

Yigitbasi, N., Gallet, M., Kondo, D., Iosup, A., & Epema, D. (2010). Analysis and modeling of time-correlated failures in large-scale distributed systems. In Grid Computing (GRID), 2010 11th IEEE/ACM International Conference on (pp. 65–72). IEEE.

Zhao, H. & Yin, Y.-l. (2011). A dynamic mechanism of risk allocation of construction project from perspective of incomplete contract theory: A theoretical model. In Industrial Engineering and Engineering Management (IE&EM), 2011 IEEE 18Th International Conference on (pp. 1816–1820). IEEE.

Zhu, Q. & Fung, R. Y. (2012). Design and analysis of optimal incentive contracts between fourth-party and third-party logistics providers. In Automation and Logistics (ICAL), 2012 IEEE International Conference on (pp. 84–89). IEEE.

Zietlow, G. (2007). Performance-based road management and maintenance contracts-worldwide experiences. In International Seminar on Road Financing and Investment (pp. 1–76).