9-2018

Methods to Support the Project Selection Problem With Non-Linear Portfolio Objectives, Time Sensitive Objectives, Time Sensitive Resource Constraints, and Modeling Inadequacies

Neil C. Ranly

METHODS TO SUPPORT THE PROJECT SELECTION PROBLEM WITH NON-LINEAR PORTFOLIO OBJECTIVES, TIME SENSITIVE OBJECTIVES, TIME SENSITIVE RESOURCE CONSTRAINTS, AND MODELING INADEQUACIES

DISSERTATION

Neil C. Ranly

AFIT-ENS-DS-18-S-040

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENS-DS-18-S-040

METHODS TO SUPPORT THE PROJECT SELECTION PROBLEM WITH NON-LINEAR PORTFOLIO OBJECTIVES, TIME SENSITIVE OBJECTIVES, TIME SENSITIVE RESOURCE CONSTRAINTS, AND MODELING INADEQUACIES

DISSERTATION

Presented to the Faculty

Department of Operational Sciences

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

Neil C. Ranly, B.S., M.S.

September 2018

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENS-DS-18-S-040

METHODS TO SUPPORT THE PROJECT SELECTION PROBLEM WITH NON-LINEAR PORTFOLIO OBJECTIVES, TIME SENSITIVE OBJECTIVES, TIME SENSITIVE RESOURCE CONSTRAINTS, AND MODELING INADEQUACIES

Neil C. Ranly, BS, MS

Committee Membership:

Jeffery D. Weir, PhD
Chair

John M. Colombi, PhD
Member

Maj Heidi M. Tucholski, PhD
Member

ADEDEJI B. BADIRU, PhD
Dean, Graduate School of Engineering and Management

AFIT-ENS-DS-18-S-040

Abstract

This dissertation presents research concerning project selection problems motivated by issues affecting Air Force information production activities. First, an optimization methodology is presented for a deterministic project selection and scheduling problem variant composed of variable intensity work activity, benefit-event deadlines, and predecessor constraints. An experiment with 1,800 problem instances is performed and the results show the methodology produces an equal or a better optimized solution compared to a methodology proposed in literature for every problem instance. Second, an optimization formulation is proposed for a project selection and sequencing problem variant including project re-execution decisions to account for the aging, information-based product of the projects. A case study finds the formulation computationally tractable and the results insightful for cost-vs-benefit analysis. Third, a technique used to address modeling inadequacies is revisited. The technique generates decision-space diverse solutions for the presentation to the decision maker. Limitations of the technique to address solution set content correlation attributes of diversity are highlighted. A technique modification is proposed to address these correlation attributes of diversity and an experiment is performed for a portfolio selection problem. Paired t-test results show the proposed technique produces significantly more diverse solution sets compared to the original technique regarding correlation-sensitive diversity measures. Fourth, a methodology is proposed to address non-constant marginal values in project selection problems' portfolio-based objectives. A branch and bound extension of an open

source non-linear programming solver is proposed and results in a reduction of the optimality gap compared to a commercial off-the-shelf solver for a problem dataset from the literature.

Acknowledgments

I would like to express my sincere appreciation to my advisor, Dr. Jeffery Weir, for his guidance and support, and my research committee members, Dr. John Colombi and Major Tucholski, for their insight and their constructive criticisms that advanced this research.

This research is in support of the Air Force, Life Cycle Management Center, Operation Research and Analysis Division (AFLCMC/OZA).

Neil C. Ranly

Table of Contents

List of Figures

Page

List of Tables

Page

METHODS TO SUPPORT THE PROJECT SELECTION PROBLEM WITH NON-LINEAR PORTFOLIO OBJECTIVES, TIME SENSITIVE OBJECTIVES, TIME SENSITIVE RESOURCE CONSTRAINTS, AND MODELING INADEQUACIES

I. Introduction

The project selection problem is a vital concern in many organizations. The crux of the problem is an organization has many project opportunities and not enough resources to pursue all the projects. The organization must choose a subset (i.e., portfolio) of all potential projects to pursue. The degree the organization pursues more valuable projects efficiently and effectively ultimately influences the long-term success of the organization.

This dissertation first provides the background and the motivating application of this research. Given the background, research questions are presented based on gaps discovered from a literature review. Next, related work is presented discovered from a review of research relating to the project selection problem and to the issues related to the problem context. Next, four papers present research regarding research question gaps, a method to address the research gaps, and results that show the method's significance. Finally, a summary provides a conclusion of the research findings and, given these findings, possible future research.

1.1 Background

A DoD intelligence project selection problem is highlighted in a report by the Government Accountability Office (Government Accountability Office, 2016). Intelligence (i.e., information) users, who include research and development organizations, system acquisition organizations, and combat organizations, request many structured intelligence products to support their missions. In the creation of an intelligence product, the "project" consumes and ties up research and developmental resources. Given a limited intelligence production resource capacity, the intelligence organizations cannot satisfy all the production requests and the defense and intelligence enterprise must select a subset of intelligence production projects to pursue given competing priorities.

Air Force acquisition organizations face the intelligence-project selection problem described above in a slightly different context, this being determining the intelligence supportability of a given system and setting their expectations to receive a set of intelligence products. For example, if the Air Force acquires a system and the intelligence is not produced to the system's expectations, this may result in an ineffective system and a waste of money. With the increasing desire to use intelligence data in sophisticated sensor and automation capabilities, acquisition organizations should consider the ability of the producers to provide the intelligence needed and set their expectations accordingly (selecting a level of expected intelligence based on a set of intelligence projects). The growing use of automation in sensor applications and decision support systems, such as the F-35, is increasing the demand for intelligence products (Government Accountability Office, 2016).

The intelligence project selection problem described above has attributes that make it complex compared to other resource allocation problems. First, the project selection problems of interest are strongly affected by events at points in time and over time, similar to issues raised in past military focused research (G. G. Brown, Dell, and Newman 2004). Budgets and personnel resource changes limit production capacity differently in distinct time periods. Also, the customers of the intelligence projects often have time requirements for use, such as an acquisition testing milestone (e.g., if the product of an intelligence project is received after a test event, there is no benefit realized from the project with regard to that test objective). These time requirements provide a deadline for benefit/objective realization.

Second, decision makers often desire a small set of diverse alternatives to manually compare and contrast as a part of a decision making process and part of the foundation of sensitivity analysis (Berntsen and Trutnevyte 2017; S.-Y. Chang, Brill, and Hopkins 1982; Hennen et al. 2017). This desire is often driven from the realization that the optimization model is incomplete based on unquantifiable aspects of the problem. For example, the nature of classified projects and an uncertain future imposes modeling limitations that an analyst cannot completely address. While a project selection problem with just 20 potential projects results in about one million possible alternatives (or 2^{20}, ignoring any possible feasibility constraints), prescribing a small, diverse set of near-optimal solutions requires a method to generate these solutions.

Third, unlike commercial profit-oriented applications that often use net present value (NPV) project consequence measurements, military applications do not readily suggest a single ratio based unit of measurement for project consequence quantification (Bullock 2006; Ewing, Tarantino, and Parnell 2006). Military applications often consider multiple objectives with different units of consequence (i.e., benefit) measurement. For example, the product of an intelligence project may ultimately save a pilot's life, may provide a vital piece of information to enable a high-risk mission to strike an adversary's center of gravity, or may just provide slightly better battlespace awareness to a pilot. The product of an intelligence project to support an acquisition or a research and development customer may give insights into the effective design of a next generation system (e.g., saving lives in future scenarios) or may provide the insight into the efficient design of a next generation system (e.g., saving money). Often measurements of these projected consequences require normalization that considers all the decision maker's fundamental objectives. In addition, the objective criteria of a project selection problem may possess non-constant marginal preferences. Traditional project selection methods do not directly support these non-linear characteristics of the portfolio-level value or utility functions (Liesiö 2014). For military acquisition decision support, Key Performance Parameters (KPPs) are one method that the defense enterprise expresses their objectives to a measurable, operational form (R. R. Hill et al. 2013). From the perspective of the intelligence project selection problem, KPPs can be viewed as a basis for holistic portfolio-level value functions.

This dissertation provides insight regarding methodologies to address these complexities observed in the intelligence-production project selection problem (or from the acquisition consumer perspective, the intelligence requirement selection problem) and develops novel methods to address the proposed research questions. In addition, this research has wider applicability given the importance of project selection and resource allocation activities as demonstrated in past research concerning crowdsourcing information production (Basu Roy et al. 2015), software engineering (W. Chen and Zhang 2013), and technology portfolio management (Dickinson, Thornton, and Graves 2001).

1.2 Research Questions

This dissertation provides methods developed to address the following questions concerning the project selection problem. These questions are concentrated on issues that researchers have not addressed completely in the literature.

1. How should organizations perform multi-objective project selection given complicating time factors, including variable-intensity-project work, resource availability changes through time and changes through time given event-benefit deadlines?

Chapter III proposes a new methodology to find the optimal solution to the project selection and scheduling problem (PSSP) with variable intensity work and deadlines. The methodology finds the problem's optimal solution for every problem, more efficiently compared to a methodology proposed in literature that fails to find the problem's optimal solution for over 1/5 of the problem instances in the experimentation dataset.

2. How should organizations perform multi-objective project selection given ageing, shareable products and project re-execution (i.e., reproduction) decisions?

Chapter IV introduces a new methodology to support solution optimization in regard to product deterioration and reproduction decisions. The chapter provides a formulation to compute a measurable value baseline with respect to time interactions on decision value evaluations. Case study results show the computational tractability and the ability of the method to produce cost-vs-benefit insight for an Air Force intelligence production planning problem.

3. How should an analyst produce a small, diverse set of project selection solutions to support decision makers applying their elusive to model knowledge to work around modeling inadequacies?
    - What quantitative measure should the analyst use to assess solution set diversity?

Chapter V develops and shows a new project selection alternative generation technique that generates significantly more decision-space-diverse alternative sets compared to an existing alternative generation technique from literature. Appendix VIII presents a new set correlation-sensitive diversity measure based on information entropy to quantify the decision-space diversity of a solution set.

4. How could organizations perform modeling and optimization of project selection value with multiple objectives holding non-constant marginal values?

Chapter VI presents progress towards an optimization method for the project selection problem with multiple, non-constant-marginal-value objectives. The chapter also presents an optimization solver for the non-linear programming formulation and shows the solver finds a better solution (reduces optimality gap) compared to a commercial-off-the-shelf non-linear solver for a problem dataset provided in literature. The proposed methodology incorporates proven multi-objective modeling techniques addressing a limitation of a recently proposed method requiring unproven techniques.

II. Literature review

The project selection problem has been researched extensively throughout the years (Kleinmuntz 2007). Klienmuntz provides a summary of popular methods often incorporated into project selection and resource allocation decision applications. In particular, he presents the following methods:

1) binary integer linear programming (BILP) formulation method for a set of projects, I, and i as a member of this set,

$$\text{maximize} \sum_{i \in I} b_i x_i$$
$$\text{subject to}$$
$$\sum_{i \in I} c_i x_i \leq \boldsymbol{c}$$
$$x_i = (0 \text{ or } 1), \forall i \in I$$

2) ranking benefit-cost ratio or profitability index method (rank and then select projects in order until resources are exhausted),

$$\text{benefit-cost ratio} \quad \frac{b_i}{c_i}$$

$$\text{profitability index} \quad \frac{b_i - c_i}{c_i}$$

3) multi-attribute value or utility method for the computation of benefits based on multiple objectives (Keeney and Raiffa 1976), including the linear-additive multi-attribute measurable value function for *n* objectives:

$$b_i = \sum_{j=1}^{n} w_j v_j (y_{ij})$$

where

x_i denotes the decision to select project i,
b_i denotes the benefits in selecting project i
c_i denotes the cost to select i
$\boldsymbol{c}$ denotes the budget
y_{ij} denotes the benefit in selecting i in regard to objective j's criteria
$v_j(\cdot)$ denotes the j^{th} objective's value function
w_j denotes the weight for the j^{th} objective

These methods are simple with regards to computing solutions, assuming the data inputs are available, though by themselves often do not address significant decision context factors. From a survey in 1975 (Baker and Freeland 1975), Baker and Freeland classify these factors and present several limitations of project-selection optimization methods:

1. Inadequate treatment of risk and uncertainty
2. Inadequate treatment of multiple criteria
3. Inadequate treatment of project interrelationships
4. No explicit use of the experience and knowledge of the SMEs
5. Inability to handle nonmonetary aspects
6. Perception that the models are difficult to understand
7. Inadequate treatment of time variations

The project selection topic, as a whole, is revisited again in 1999 (Heidenberger and Stummer 1999). Heidenberger & Stummer review project selection and resource allocation quantitative modeling approaches and organize their findings as depicted in Figure 1. Heidenberger and Stummer (1999) present a few more topics in relation to project selection, including the use of simulation models in project selection activities. G. G. Brown et al. (2004) review the history of mathematical programming in military capital planning activities. They highlight efforts that incorporated time dependencies, interactions between decisions, synergistic effects among decision variables, and aged inventory issues into mathematical programming models.

To support this research, the issues of project selection problems and associated solution techniques are classified into the following categories: benefit measurement; project completion attributes; optimization method; and addressing optimization-method inadequacies. Sub-issues are italicized. Associations to the subsequent chapters are presented as appropriate. Following discussion of these main issues, Table 1 presents an organization of references linked to important themes.

- Benefit measurement methods
  - Comparative Models
    - Q-sort approach
    - Analytical hierarchy process
  - Scoring Approaches
    - Checklist approach
    - Traditional scoring models
    - Multiattribute utility analysis
  - Traditional Economic Models
    - Economic indexes
    - Discounted cash-flow methods
    - Options approach
  - Group Decision Techniques
- Decision and Game Theory
  - Decision-Tree Approaches
  - Game-theoretical Approaches
- Mathematical Programming
  - Linear Programming Models
  - Non-linear Programming Models
  - Integer Programming Models
  - Goal Programming Models
  - Dynamic Programming Models
  - Stochastic Programming Models
  - Fuzzy Mathematical Programming Models
- Simulation Models
- Heuristics
- Cognitive Emulation
  - Statistical Approaches
  - Expert Systems
  - Decision Process Analysis

FIGURE 1. AN ORGANIZATION OF PROJECT SELECTION TOPICS (HEIDENBERGER AND STUMMER 1999)

2.1 Benefit measurement

Optimizing project selection, whether concerning a single or multi-objective, relies upon benefit measurements. For example Heidenberger and Stummer (1999) presents a listing of techniques to measure benefits (Figure 1). Regarding the *perspective of measurement*, most research employs a project centric approach, which this dissertation defines as the measurement assumption that the value a project adds to all possible portfolios does not depend on what other projects are in the portfolio. For example, summing the selected projects' scores to compute the overall portfolio's value or ranking the benefit-vs-cost value and selecting the projects in order until resources are exhausted employ the project-centric perspective assumption. Another perspective is a holistic approach enabling non-constant marginal value preferences to be measured (i.e., the value of a project depends on what other projects are in the portfolio) and explicitly addressed (Liesiö 2014).

Regarding the *dimensionality of measurement*, the project selection problem may be formed for a single objective, while other problems may be formed for multiple fundamental objectives (Ghorbani and Rabbani 2009; Golabi, Kirkwood, and Sicherman 1981; Medaglia, Graves, and Ringuest 2007; Rabbani, Aramoon Bajestani, and Baharian Khoshkhou 2010). Researchers suggest different *foundations of objective measurement*. For instance researchers propose benefit measurements using:

- natural ratio units, such as net present value (Dickinson, Thornton, and Graves 2001)
- relative comparisons, such as AHP (Amiri 2010)
- normative techniques, such as measurable value functions (Golabi, Kirkwood, and Sicherman 1981; Liesiö and Punkka 2014)
- risk & uncertainty (i.e., utility) approaches (Liesiö and Salo 2012)

An additional aspect of measurement includes *project set interaction* effects of benefits. For example, two projects may display a synergistic or antagonistic effect when both are chosen in a solution (Stummer and Heidenberger 2003). Another important aspect of is the *temporal interactions* on benefits (Stummer and Heidenberger 2003). This dissertation distinguishes between two types of temporal benefit interactions:

deterioration such as discounting and appreciation, and *event-based* such as benefit deadlines (J. Chen and Askin 2009).

This dissertation presents two novel, project-centric perspective methods. Chapter III presents a method to address multiple event-based, benefit deadlines. Chapter IV presents a method to address deterioration and event-based benefit measurement issues. Chapter VI introduces a holistic-perspective benefit-measurement-method incorporating a normative measurement foundation of a multi-objective project selection problem.

2.2 Project completion attributes

Project selection problems possess issues that restrict when, how, with what, by who, and where the projects are completed (Fox, Baker, and Bryant 1984; Zhao and Huang 2017). Similar to benefit measurements, optimizing project selection relies upon project *cost measurements* (i.e., resource requirement measurements). Important aspects of cost measurement include the *dimensionality of measurement* (resource types and the nature of resource types needed for project completion, such as renewable vs consumable resources) and *project set interactions* effects regarding costs (Zhao and Huang 2017).

Additional project completion issues may arise from temporal attributes of scheduling projects. Research involving resource constrained project scheduling problem (RCPSP) provides a foundation for addressing these issues (Hartmann and Briskorn 2010). Temporal issues include *temporal order constraints* (i.e., predecessor constraints) (Tofighian and Naderi 2015), *temporal completion options,* which includes *intensity attributes* (X. Li et al. 2015), such as whether projects are of fixed duration or of variable intensity (Askin 2003; J. Chen 2005), and *renewable resources assignment options.* For example, *renewable resources assignment options* may assume a resource entity is assigned to a single project (Taylor, Moore, and Clayton 1982) or to multiple projects over the course of the planning horizon and may assume a resource entity's availability changes over time. Another aspect of project completion attributes is the *decision nature of the resource constraints* (i.e., fixed or variable budget).

Chapter III presents a project selection and scheduling method to address variable *intensity attributes*, *temporal order constraints*, and *resource availabilities changes* in time. Chapter IV presents a project selection and scheduling method to address *temporal order constraints* derived from re-execution decisions. Chapter III and Chapter IV demonstrate an approach to the *decision nature of the resource constraints* issue given the relatively quick optimization computation times enabled by each project selection and scheduling optimization methodology.

2.3 Optimization method

Project selection research suggests numerous methods to generate an optimal or a near-optimal solution. Mathematical programming models are a common method with deterministic input parameter assumptions (Kleinmuntz 2007). The literature proposes genetic algorithms, simulated annealing, and other heuristic search optimization methods to address stochastic input parameter assumptions (Bhattacharyya, Kumar, and Kar 2011; Crama and Schyns 2001; Huang 2007). Project selection and scheduling research suggest ant colony optimization (Tofighian and Naderi 2015) and genetic algorithms (Rabbani, Aramoon Bajestani, and Baharian Khoshkhou 2010) methods.

Chapter III and Chapter IV proposes an optimization method for project selection and scheduling problem deterministic variants utilizing a computationally tractable binary linear programming formulation. Chapter VI proposes an optimization method for a multi-objective project selection problem with non-constant marginal values and deterministic inputs using a branch and bound enumeration technique over a computationally quick non-linear programming solver.

2.4 Addressing optimization modeling inadequacies

Optimization methods for problem selection problems rely on assumptions. Researchers suggest extra caution in outright acceptance of optimization solutions considering potential optimization modeling errors due to invalid assumptions which should be addressable by the decision maker after the optimization method provides a solution. Research suggests the use of *decision support techniques*. These techniques include the presentation of a decision-space-diverse set of alternatives (E Downey Brill, Chang, and Hopkins 1982; DeCarolis et al. 2016), the presentation of a Pareto optimal set of alternatives (Kangaspunta, Liesiö, and Salo 2012), an interactive-iterative refinement of modeling expressiveness (Argyris, Figueira, and Morton 2011; Nowak 2013), and the discovery of robust elements of optimal portfolios over an incomplete information parameter space (Fliedner and Liesiö 2016; Liesiö, Mild, and Salo 2008). Chapter V presents research into an alternative generation technique that generates a decision-space diverse set of project selection solutions.

2.5 Cross Reference Table

To summarize the background material related to the research questions, the literature is categorized into the following themes (Table 1):

- multi-objective methods
- measurable value functions
- uncertainty & utility functions
- preference elicitation methods
- optimization methods
- approaches to non-linear issues
- generating diverse alternatives
- incomplete information
- time related issues
- information production planning topics

TABLE 1. CROSS REFERENCE TABLE OF INTEREST THEMES TO REFERENCES

References	Multi-objective Methods	Measurable Value Functions	Uncertainty & Utility Functions	Preference Elicitation Methods	Optimization Methods	Approaches to Non-linear Issues	Sensitivity Analysis	Generating Diverse Alternatives	Incomplete Information	Time related issues	Information Production Planning
(Abbas 2003)	x		x						x		
(Abbas and Howard 2005)	x		x	x		x					
(Ahn 2011)	x			x							
(Amiri 2010)	x			x					x		
(Amorim, Günther, and Almada-Lobo 2012)										x	
(Archer and Ghasemzadeh 1999)							x				
(Argyris, Figueira, and Morton 2011)	x			x	x	x	x				
(Askin 2003)					x					x	
(Baker & Freeland, 1975)						*survey*					
(Basu Roy et al. 2015)										x	x
(Baugh Jr., Caldwell, and Brill Jr. 1997)							x	x			
(Beaujon, Marin, and McDonald 2001)						x					
(Bessai and Charoy 2017)					x					x	x
(Beynon and Curry 2000)	x			x							
(Bickel and Smith 2006)					x					x	
(Bilbao et al. 2007)				x							
(Bhattacharyya, Kumar, and Kar 2011)	x				x	x				x	
(Blecic, Cecchini, and Trunfio 2012)			x		x	x				x	
(Bobbio et al. 2001)	x		x								
(Bonini 1975)			x								
(E Downey Brill, Chang, and Hopkins 1982)								x			
(E.D. Brill et al. 1990)								x			

References	Multi-objective Methods	Measurable Value Functions	Uncertainty & Utility Functions	Preference Elicitation Methods	Optimization Methods	Approaches to Non-linear Issues	Sensitivity Analysis	Generating Diverse Alternatives	Incomplete Information	Time related issues	Information Production Planning
(G. G. Brown, Dell, and Newman 2004)			x		x	x				x	
(D. B. Brown and Smith 2013)					x					x	
(Brucker et al. 1999)										x	
(Carazo et al. 2010)	x				x	x				x	
(Čepin and Mavko 2002)			x							x	
(D.-Y. Chang 1996)	x										
(S.-Y. Chang, Brill, and Hopkins 1982)								x			
(Charnes, Cooper, and Rhodes 1978)	x						x				
(J. Chen 2005)					x	x				x	
(J. Chen and Askin 2009)	x				x					x	
(W. Chen and Zhang 2013)	x				x	x				x	
(Cochran, Horng, and Fowler 2003)	x									x	
(Coffin and Taylor 1996)					x					x	
(Crama and Schyns 2001)					x	x					
(Danielson et al. 2007)	x		x	x			x		x		
(Danna and Woodruff 2009)					x			x			
(DeCarolis et al. 2016)								x			
(Dickinson, Thornton, and Graves 2001)	x		x	x	x	x	x			x	
(Distefano and Puliafito 2009)	x										
(Doemer et al. 2004)	x				x	x		x		x	
(Drineas et al. 2004)								x			
(Durga Rao et al. 2009)			x								
(Dyer and Sarin 1979)		x	x								
(Dyer 1990)		x	x								
(Eilat, Golany, and Shtub 2006)	x				x		x				
(Erkut, ReVelle, and Ülküsal 1996)								x			
(Ewing, Tarantino, and Parnell 2006)	x	x		x							
(Fliedner and Liesiö 2016)	x				x		x		x		
(Fox, Baker, and Bryant 1984)					x	x					
(Ghorbani and Rabbani 2009)	x				x	x				x	
(Golabi, Kirkwood, and Sicherman 1981)	x	x		x	x		x				
(Graves, Ringuest, and Bard 1992)								x			
(Gren 2017)										x	
(Gröwe-Kuska, Heitsch, and Römisch 2003)			x			x					
(Guikema and Milke 2003)	x						x	x			
(Gustafsson and Salo 2005)	x				x					x	
(Gutjahr et al. 2010)	x				x	x				x	
(Hartmann and Briskorn 2010)										x	
(Hassanzadeh et al. 2014)					x					x	
(Heidenberger and Stummer 1999)						*survey*					
(Hennen et al. 2017)							x	x			
(Hess 1962)					x					x	
(Henriksen and Traynor 1999)	x			x							
(R. R. Hill et al. 2013)			x								
(Hu et al. 2008)	x					x					
(Huang 2007)	x			x	x	x			x		
(Huang 2008)			x								
(Huynh and Simon 2016)				x					x		
(Jiang et al. 2011)	x		x	x							

References	Multi-objective Methods	Measurable Value Functions	Uncertainty & Utility Functions	Preference Elicitation Methods	Optimization Methods	Approaches to Non-linear Issues	Sensitivity Analysis	Generating Diverse Alternatives	Incomplete Information	Time related issues	Information Production Planning
(Kangaspunta, Liesiö, and Salo 2012)	x				x		x				
(Keeney 1981)		x				x					
(Keeney 2002)	x	x	x	x							
(Keisler 2004)					x		x				
(Khorramshahgol and Gousty 1986)				x	x						
(Kopanos, Kyriakidis, and Georgiadis 2014)										x	
(Kuei H. 1994)	x				x	x					
(Lee and Kim 2001)	x			x							
(X. Li et al. 2015)					x	x				x	
(Liang and Li 2008)	x					x					
(Liesiö, Mild, and Salo 2007)	x				x		x				
(Liesiö, Mild, and Salo 2008)	x				x	x	x		x		
(Liesiö and Salo 2012)	x		x				x				
(Liesiö and Punkka 2014)	x	x	x				x				
(Liesiö 2014)	x	x	x	x	x	x					
(Liu and Wang 2011)					x	x				x	
(Loch and Kavadias 2002)										x	
(Lucas et al. 2017)								x			
(Luce and Tukey 1964)		x									
(Maheswari and Varghese 2005)										x	
(D Makowski et al. 2000)							x	x			
(David Makowski et al. 2001)							x	x			
(Malakooti 1991)		x		x							
(Mandakovic and Souder 1985)					x					x	
(Mavridis, Gross-Amblard, and Miklós 2016)											x
(Meade and Presley 2002)	x					x					
(Medaglia, Graves, and Ringuest 2007)	x		x		x	x			x		
(Medaglia et al. 2008)	x				x					x	
(Montiel and Bickel 2014)			x	x			x				
(Morse 1980)								x			
(Naber and Kolisch 2014)										x	
(Nemhauser and Ullmann 1969)					x						
(Nowak 2013)							x	x			
(Pape 2017)				x					x		
(Pahl and Voß 2014)										x	
(Petit 2012)			x								
(Phillips and Bana e Costa 2007)	x			x							
(Piper and Zoltners 1976)							x	x			
(Preacher, Curran, and Bauer 2006)						x					
(Rabbani, Aramoon Bajestani, and Baharian Khoshkhou 2010)	x				x	x					
(Rahman et al. 2015)					x					x	x
(Razmi and Rafiei 2010)	x				x	x					
(Ringuest and Graves 1990)	x		x							x	
(Ringuest, Graves, and Case 2004)			x			x					
(Rong, Akkerman, and Grunow 2011)					x					x	
(Band and Andrews 2004)			x								
(A. A. Salo and Hämäläinen 2001)	x	x		x					x		
(A. Salo and Punkka 2005)	x			x			x		x		

References	Multi-objective Methods	Measurable Value Functions	Uncertainty & Utility Functions	Preference Elicitation Methods	Optimization Methods	Approaches to Non-linear Issues	Sensitivity Analysis	Generating Diverse Alternatives	Incomplete Information	Time related issues	Information Production Planning
(Saltelli 2002)							x				
(Santhanam and Kyparisis 1996)					x	x					
(Sarin 1982)		x	x								
(Sayın 2000)								x			
(D. A. Scott et al. 2009)				x							
(D. Scott and Suppes 1958)		x									
(Sefair and Medaglia 2005)	x									x	
(Shariatmadari et al. 2017)					x					x	
(Sinha, Majumder, and Manjunath 2016)										x	x
(Sobol and Kucherenko 2005)			x				x				
(Spetzler and Staël Von Holstein 1975)			x				x				
(Srivastava, Connolly, and Beach 1995)	x			x							
(Steuer and Harris 1980)								x			
(Stummer and Heidenberger 2003)	x					x				x	
(Stummer, Kiesling, and Gutjahr 2009)	x				x					x	
(Sun and Ma 2005)					x					x	
(Tavana, Abtahi, and Khalili-damghani 2014)										x	
(Taylor, Moore, and Clayton 1982)						x					
(Tofighian and Naderi 2015)	x				x					x	
(Toppila and Salo 2017)	x				x				x		
(Trainor, Parnell, and Kwinn 2004)	x	x		x							
(Tran-Thanh et al. 2014)										x	x
(Udell and Boyd 2013)					x	x					
(Ulrich, Bader, and Thiele 2010)					x			x			
(Vilkkumaa, Salo, and Liesiö 2014)	x			x					x		
(Villarreal and Karwan 1981)	x				x						
(Voll et al. 2015)								x	x		
(Wall 1996)										x	
(Walls, Morahan, and Dyer 1995)				x		x					
(Kleinmuntz 2007)						*survey*					
(Wilck et al. 2016)	x				x	x	x				
(Wilbaut, Hanafi, and Salhi 2007)	x				x						
(Winterfeldt and Edwards 1986)	x	x	x	x							
(You and Yamada 2007)					x						
(Zan et al. 2013)			x				x				
(Zechman and Ranjithan 2007)								x			
(Zhao and Huang 2017)					x					x	
(Zuluaga, Sefair, and Medaglia 2007)					x	x				x	

2.6 Literature Summary

In total, this dissertation considers over 150 references published from 1957 to 2017 relating to project selection problems and issues pertinent to the research questions.

By most referenced research theme, the dissertation references approximately 61 references concerning multi-objective issues, followed by 58 references concerning optimization methods, and 50 references researching time related issues. The individual chapters present additional literature as appropriate and provide a detailed discussion of the literature with respect to the research questions.

III. Project selection and scheduling with variable intensity work

Variable-intensity work activities require a fixed amount of total time commitment from renewable resource(s) with the option to apply the work in inconsistent amounts over time. For a project selection and scheduling problem composed of variable-intensity work, this paper proposes a methodology to optimize the value of a portfolio of projects with regards to deadline dependent benefits. Other attributes of the problem include project-predecessor dependencies and varying levels of resource availability over time. The methodology only requires the discretization of the planning horizon based on benefit deadlines and reduces the number of linear programming decision variables in an integer programming model compared to other proposed formulations. Considering a dataset consisting of 1,800 problem instances, the results show the methodology results in significantly better solutions and more quickly computes optimal solutions compared to a previously suggested variable-intensity project selection and scheduling optimization methodology from literature.

3.1 Introduction

Organizations often encounter situations involving resource limitations that restrict the pursuit of all conceived projects. This forces organizations to select a subset of the projects that maximizes the organizations' objective(s) (Kleinmuntz 2007). In more complex situations, organizations must perform the selection and plan the execution of projects with respect to time dependent objectives specified by a deadline, such as a fixed time event desiring the usage of the project's product to realize additional benefits, and with respect to resource availability changes, such as employee leave absences, vacations, contractual short-term workforce support, resource maintenance activities, and training activities. In even more complex situations, some projects need to be completed before other projects can start.

This paper proposes a methodology to address a variant of this problem. In this variant, symmetrical skilled resource entities, such as cross-trained analysts or engineers, can work together to finish a project activity quicker in a linear manner, which Chen (2005) refers to as variable-intensity work. For example, an activity that requires one month of work can be completed by one resource entity in one month, completed by two resource entities in 2 weeks, or completed by three resource entities with 2 of these entities contributing one week of work and the remaining entity contributing two weeks of work. Likewise, more resource entities could complete the work of a project in a shorter time. In this variant, the linear resource-work-intensity to duration relationship holds up to the number of resource entities available at any given time and the intensity is permitted to vary throughout the life of the project. This results in project activities having a variable duration based on the intensity of the work provided by the assigned resources to complete the project. Also, the problem consists of projects that are simple to start resulting in zero startup costs. Organizations may encounter this problem in production environments with flexible processing agents and with teams that cultivate variable-intensity options to reduce project duration in ad-hoc time-critical situations where every resource entity possesses the skills needed for all potential project-work to support flexible, robust operations.

Considering an 1,800 problem instance dataset, the proposed methodology and a methodology from literature are used to solve each problem instance. The results show the proposed methodology demonstrates positive properties with regard to computational tractability and results in significantly better solutions when compared to solutions from a method from literature that overly restricts solutions to the discretization of the planning horizon. The methodology employs a binary integer linear programming model that incorporates fewer decision variables. By decomposing the problem into two stages, the proposed method only discretizes the planning horizon using the benefit deadlines removing the need for a user to choose a step size for resource assignment in the optimization model. The methodology finds an optimal schedule on a continuous time dimension. The previously proposed mixed integer linear programming method fails to find optimal solutions in over 20% of the problem instances due to overly restricting solutions to discrete time model formulations.

Next, the paper presents findings from a literature review related to project selection methods that consider the execution of projects over time. Section 3.3 presents the optimization methodology and how the methodology decomposes the problem into two stages. Then, Section 3.4 presents the details of an experiment comparing the proposed methodology to a methodology proposed in literature. Section 3.5 presents the results, a discussion of the results, and possible extensions to the methodology. The paper concludes with some closing remarks regarding this research and possible future research topics.

3.2 Literature Review

Researchers have studied the project selection problem, also known as the resource allocation problem and as the project portfolio selection problem, for over forty years (Baker and Freeland 1975; Heidenberger and Stummer 1999). In 2007, Klienmuntz (2007) provides a review of project selection research and popular techniques applied to the problem. More recently, Liesiö (2014) and Liesiö & Punkka (2014) revisit and propose extensions to measurable multi-attribute benefit measurement techniques in regard to project selection problems, and Fliedner & Liesiö (2016) and Toppila & Salo (2017) study techniques to improve the efficiency and effectiveness of finding problem selection problem solutions with incomplete information.

Before the turn of the century, little research existed considering the simultaneous project selection and scheduling problem (PSSP). In 1962, Hess (1962) proposes a dynamic programming method to support project selection with time affected decision factors. In 1982, Taylor III, Moore, and Clayton (1982) employ non-linear integer goal programming to support allocation of individual researcher resources to projects with a solution of static researcher assignments through time. In 1996, Coffin and Taylor (1996) propose a beam search heuristic method with a scheduling heuristic to solve a project selection and scheduling problem for projects with fixed durations decomposed into 3 stages.

After the turn of the century, research regarding the simultaneous PSSP intensified and continues with recent research with projects of fixed durations (Shariatmadari et al. 2017). The first proposed method discovered to support variable-intensity work in a project selection and scheduling problem is from (Askin 2003). Askin presents a formulation of the problem and provides a heuristic based method to generate solutions. Chen (2005) revisits project selection and scheduling with variable-intensity tasks and provides a mixed integer linear programming (MILP) model to support solution optimization. The formulation permits the bounding of intensity level of project work activities and multiple resource types. Other types of variable intensity work research exists. Kolisch & Meyer (2006) leverage a genetic algorithm based method to support multi-mode project selection and scheduling. Multi-mode scheduling employs a discrete set of variable-intensity levels. Li, Fang, Guo, Deng, and Qi (2015) propose a solution methodology to project selection problems with project "divisibility". "Divisibility" is a type of variable-intensity work that considers solutions that reduce the normal intensity work activity of a project by spreading the work into different time periods. They provide a MILP optimization method that allocates resources to discrete time periods to fund the execution of the projects.

In order to optimize the benefit of project portfolios, the measurement of benefits is a critical issue in project selection problems (Baker and Freeland 1975; Heidenberger and Stummer 1999; Kleinmuntz 2007). The problem of interest assumes the benefits from completing each project in regard to an event deadline are measured with net-present value (NPV) measurements or multi-attribute value functions (Golabi, Kirkwood, and Sicherman 1981). The incorporation of multiple, time dependent benefit effects in an optimization method is another issue. The problem of interest derives benefits from the completion of projects before deadlines and captures the benefits at the deadlines. Numerous efforts suggest measuring the possible benefits at discrete points in time and encoding the benefits and the project work decisions through time at these discrete points in an optimization model (Askin 2003; Bhattacharyya, Kumar, and Kar 2011; Carazo et al. 2010; J. Chen 2005; Dickinson, Thornton, and Graves 2001; Medaglia et al. 2008; Sefair and Medaglia 2005; Tofighian and Naderi 2015; Zuluaga, Sefair, and Medaglia 2007). Chen and Askin (2009) present a model and method to solve a project selection and fixed intensity task scheduling problem (PSFITS) with time dependent returns. Sun & Ma (2005) propose an iterative integer linear programming (ILP) model method over discrete time periods to find solutions to the project selection and scheduling problem with fixed length activity durations. Medaglia et al. (2008) present the application of a multi-objective, discrete-time-measured project selection and scheduling problem with a MILP model based solution. Sefair and Medaglia (2005) employ a dual objective optimization to also minimize the NPV variability through time. Zuluaga et al. (2007) incorporate NPV project interactions effects in time. Bhattacharyya et al. (2011) provide a fuzzy formulation of direct and project interactions benefits in time and proposed a multi-objective genetic algorithm solve method. Ghorbani & Rabbani (2009) and Tofighian & Naderi (2015) research the ability to optimize a time-dependent project selection and scheduling problem while minimizing the changes in resource utilization between time periods. Similar to these methods to represent benefits in time, the proposed methodology discretizes the planning horizon and measures the potential benefits in time at these points. Unlike these methods where a user must make a decision on discretization time step size, the proposed methodology only needs the unique deadline times to discretize the planning horizon.

The problem of interest includes technical interdependencies that require an activity's predecessors to be completed before the start of the activity. The incorporation of project interactions has been researched from numerous perspectives and applied to nuanced PSSP attributes. Zuluaga et al. (2007) present a method to solve project selection and scheduling problems with technical, resource, and benefit interdependencies. Li et al. (2015) develop and evaluate a project selection model supporting multi-time periods, a single objective (NPV), dependent projects, and project interdependences. Zhao & Huang (2017) propose a methodology to consider projects' non-renewable and renewable resource interactions. They propose a generic algorithm to solve the non-linear formulation. Dickinson et al. (2001) share research into the application of a non-linear, discrete multi-time period optimization model and business process to support the selection and scheduling of fixed duration projects with dependency interactions encoded as a dependency matrix. Liu & Wang (2011) present a model for project and scheduling problems with multiple time dependent resource constraints. The proposed methodology models the technical predecessor interdependencies slightly differently by allowing predecessors to be completed in the same time period in a mathematical model and post-processing the mathematical model optimal solution with the flexibility of variable intensity work to generate a schedule that respects the predecessor constraints.

Part of the PSSP problem of interest is assigning the project work to renewable resources (e.g., flexible machines, analysts, or engineers) explicitly and respecting the resources' work availability constraints. Researchers have studied project selection with regard to employee scheduling requirements and other time dependent resource issues. Gutjahr, Katzensteiner, Reiter, Stummer, & Denk (2010) propose a formulation to a project selection and employee assignment problem over multi-time periods and compare two meta-heuristic solution techniques: a genetics algorithm technique (NSGA-II); and an ant colony optimization technique (P-ACO). Stummer, Kiesling, & Gutjahr (2009) present a decision support system method for multi-criteria project selection and employee (with different competencies) scheduling. The former methodologies, while accounting for employee competences, do not address projects realizing benefits at time dependent deadlines of the PSSP problem of interest. They also assume the employees' availability is static throughout the planning horizon (i.e., a fixed number of employees throughout the planning horizon that do not change in time due to hiring additions, temporary work, or time-off). The methodology proposed permits resource (e.g., employee) availability changes in time.

Numerous research exists regarding the resource-constrained project scheduling problem (RCPSP). Brucker, Drexl, Möhring, Neumann, & Pesch (1999) suggest a standard notation, classification for RCPSP. They also review the methods to produce solutions for the problem. Hartmann & Briskorn (2010) review the literature regarding the RCPSP. W. Chen & Zhang (2013) develop a method to find solutions to the software task scheduling using Ant Colony optimization. Kopanos et al. (2014) provide a new formulation for the RCPSP and compare it extensively to a number of other methods. Numerous researchers have proposed methods to address variable-intensity work activities in the RCPSP (Kogan and Shtub 1999; Leachman, Dtncerler, and Kim 1990; Węglarz 1981). Naber and Kolisch (2014) study the effectiveness and efficiency of various mixed integer programming models for the project scheduling problem with flexible resource profiles. They find favorable computational properties in a variable-intensity model formulation. Kopanos, Kyriakidis, and Georgiadis (2014) propose a method to support continuous time scheduling.

A limitation of the discovered variable intensity PSSP methods discussed above is that scheduling imposed requirement to discretize the planning horizon. All the MILP techniques reviewed relied on the time discretization to assign resources to projects by discrete time periods and ensure project predecessors are completed in preceding discrete time periods. For example, to make resource allocations that could change day-by-day, the reviewed techniques require the planning horizon to be discretized by days. Choosing a too fine discretization may cause a computationally intractable problem and/or the reliance on more search iterations. Choosing too large of a discretization limits the solver from finding optimal solutions that rely upon completing projects within the same time period. The methodology proposed in the next section removes the resource-in-time allocation sensitivity of this requirement for the PSSP variant of interest while retaining positive computational properties of a small binary linear programming model.

In consideration of the methodology in regard to the application, the problem assumes variable-intensity work scales up to the number of resource entities available at any given time. Treating work as variable-intensity requires caution due to the potential work inefficiencies at high levels of intensity (Gren 2017); the result of this possible ignorance is referred to as the “Mythical Man Month” (Brooks 1995). Research exists addressing the significance of this assumption in various project management situations (Hsia, Hsu, and Kung 1999; Williams, Shukla, and Antón 2004). This research demonstrates a method to analyze project selection and scheduling results to support pre-decision discussions with the decision maker and the stakeholders to review a solution's reliance on high intensity work.

3.3 Methodology

First, note without loss of problem applicability, the methodology does not make a distinction between a project, activities, and tasks as expressed in RCPSP literature. The methodology references all workable entities as projects. A RCPSP task or activity in the formulation below is denoted as a project without any benefits but required for another task to start or for a project benefit to be realized. By only using one level of work categorization, the methodology can find solutions employing any subset of work units (i.e., the methodology does not artificially constraint the selection and scheduling of work units only if a larger set of work units is selected as a whole).

The formulation below assumes all projects are optional. Section 3.5.1 provides an extension that enables the enforcement of mandatory projects.

TABLE 2. SUMMARY OF NOTATION

Sets	
D	set of events (with a deadline) in which benefits are realized, where d denotes an element of the set and t_d denotes the deadline for event d
T	set of unique deadline times $t_d, \forall d \in D$
P	set of selectable projects, where p denotes an element of the set
E	set of predecessor constraints, i.e., $\{(j,k)$: project j must be completed before project k starts$, j, k \in P\}$
R	set of symmetric skilled project execution resource entities, such as engineers, where r denotes an element of the set
W_r	set of time windows tuples $(w_{\text{start}}, w_{\text{end}})$ that resource entity r is available to be assigned a project to work, w_{start} representing the start and w_{end} representing the end of the time window

Parameters	
$v_d(p,t)$	net present value function or normalized measurable value function if project p is completed by time t in regard to event d (the function returns 0 if $t > t_d$)
c_p	variable intensity time required (the cost) to complete project p
$\gamma_{t'}$	number of renewable resource units available up to time t'
Decision variables	
x_{pt}	binary decision to complete project p by time t
Auxiliary Functions (used in stage 2's scheduling algorithm)	
$c_{\text{total}}(R',p)$	computes the total amount of work for project p from resource set R'
$t_{\text{predecessor finish}}(p)$	rinds the latest time a predecessor of p is currently scheduled to be worked
$c_{\text{reallocate}}(j,r',\delta)$	partitions and adjusts work on project j after δ to resource r' in order to complete project j sooner so that resources' working j after δ will have availability time that equals resource r' availability time (if possible) and returns the amount of work time adjusted
$c_{\text{free}}(R',t)$	computes total uncommitted work time held by resources in R' up to time t

The variable-intensity property of the project work allows a decomposition of the problem into two stages. In the first stage, the methodology models part of the problem as binary linear programming model (BILP), only discretizing the time planning horizon by the significant changes to the projects' benefits with respect to the overall effort's deadline-dependent objectives. The resulting optimal solution indicates the projects selected and outlines the schedule by project completion requirements (i.e., sequence). Stage one of the methodology considers the overall resource needs and not the specific start and end times for each project. Theorem 1 states a feasible schedule exists given a BILP model solution. Stage one's BILP formulation is presented below.

$$\max \sum_{p\in P}\sum_{t\in T}\sum_{d\in D} v_d(p,t)x_{pt} \tag{1}$$

Subject to:

$$\sum_{p \in P} c_p \left(\sum_{t \le t'} x_{pt} \right) \le \gamma_{t'}, \forall t' \in T \tag{2}$$

$$\sum_{t \in T} x_{pt} \le 1, \forall p \in P \tag{3}$$

$$x_{kt'} - \left(\sum_{t \le t'} x_{jt} \right) \le 0, \forall (j,k) \in E, \forall t' \in T \tag{4}$$

$$x_{pt} \in \{0,1\}, \forall p \in P, \forall t \in T \tag{5}$$

Objective (1) maximizes the project portfolio's total value. The resource time windows are condensed to a sum of resource units available up to the end of each deadline-distinguished time period t', denoted $\gamma_{t'}$ in constraint (2) and computed using equation (6). Notice the time windows for each resource can start and stop anywhere; the methodology does not require them to align with the event-deadline-based discretization strategy. The second stage explicitly assigns work given the time window restrictions. Constraint set (2) ensures that the total work units needed by all the projects selected to be finished up to the end of each deadline-defined time period do not exceed the total number of resource units made available up to the end of the time period. Constraint set (3) ensures all projects are selected-to-be-completed no more than once over the time horizon. Constraint set (4) enforces the predecessor requirements. Constraint set (5) denotes the binary nature of the decision variables.

$$\gamma_{t'} = \sum_{r \in R} \sum_{w \in W_r} \max\{(\min\{t', w_{\text{end}}\} - w_{\text{start}}), 0\} \tag{6}$$

The BILP leverages and reflects the flexible nature of variable-intensity projects. Specifically, a project's predecessors can be completed in the same (or earlier) discretized time period if enough resource units are available up to that time. The proof that this is permissible relies on the variable-intensity attribute of the projects' work.

Theorem 1. There exists a schedule over a given time period for a set of selected projects composed of variable-intensity work attributes if the set of projects include every project members' predecessor(s) and the total amount of resource units available over the time period meet or exceed the sum of the selected projects' work requirements.

Proof. First given any set of selected projects $P_{s,t'} \coloneqq \{p: x_{pt} = 1, t = t' \} \subset P$ containing every predecessor of each project in the set, there exists an ordered list of the projects such that each project in the list only requires projects preceding itself in the list; if an ordered list does not exist then there is a predecessor cycle defined in the predecessor graph which by definition is not allowed. By definition of variable-intensity work, a scheduling algorithm can take this ordering and complete the projects in sequence at the highest intensity available meeting the predecessor requirements. If $P_{s,t'}$ also results in $\sum_{p \in P_{s,t'}} c_p \leq \gamma_{t'}$ (i.e., the total resource units available for work met or exceeds the amount of work required), the selected projects will be completed at or before the end of the time period. For example, if enough resource time-units in a given time period are available for project p_A and project p_B and project p_A is a predecessor of project p_B, then the variable-intensity property permits first completing project p_A using the maximum intensity so that the remaining resource time-units available after the competition of project p_A satisfy project p_B's work needs.

Allowing extra unallocated resources to start projects scheduled to be completed by the next time period and Theorem 1, a feasible schedule exists for each subsequent time period through the whole planning horizon. Thus, a feasible schedule exists for stage one's optimal BILP solution.

Stage two takes the solution to the BILP defined above and builds the detailed schedule by adjusting the intensity on a continuous timeline to meet the predecessor constraints and completion requirements. It executes an iterative resource assignment algorithm to allocate resources from the start of the planning horizon to the end of the planning horizon. See Figure 2 for the algorithm pseudo code. This is just one potential algorithm to complete this stage; other algorithms incorporating other heuristics to fulfill this stage that may be better suited to other applications (e.g., such as employing paired assignments when possible to support the software engineering practice of pairing programmers together to complete coding tasks (Williams, Shukla, and Antón 2004)).

The algorithm presented in Figure 2 assigns work to the first available resource unless a predecessor project assignment's resource(s) can complete the project before the end of the active time period. It employs high intensity work assignments (i.e., assigning a project to more than one resource at a time) when a project cannot be completed by a single resource before the completion requirement derived from stage one and with respect to predecessor constraints. It attempts to minimize high intensity work by first allocating the project work to all of a resource's work-availability-time as needed and then adjusting to account for any successor projects afterwards. If a successor project causes a lock situation (i.e., the algorithm is unable to adjust and move work from one resource to a resource that has unallocated availability earlier than the former in order to have the resources have availability concurrently without violating a predecessor constraint), all the assignments after the first available resource's availability time within in the resource context group, denoted R', are reallocated according to the projects' predecessor constrained order using variable-intensity work to the level specified by the cardinality of the resource context group through time as the proof of Theorem 1 relies upon. Note that the project resource assignments are made in succession. For example, if one project uses 0.5 time work units of a resource, the next project to be assigned to this resource starts right after the previous project assignment. This results in a continuous time schedule that generally avoids work activity preemption.

inputs: unique deadlines T, resource set R, stage one's optimal solution (project completion requirements x_{pt})

for each t **in** T *(and* T *is ordered earliest to latest)*

- $P_t \leftarrow \{p: x_{pt} = 1\}$
- $P_s \leftarrow$ elements of P_t sorted by predecessor dependencies
  projects with predecessor(s) in P_t *come after all of its predecessor(s)*
- **for each** p **in** P_s
  - $R' \leftarrow \emptyset$
  - $r' \leftarrow$ resource in R that has the earliest free time available for work allocation
  - $\delta \leftarrow$ first time that resource r' is free
  - **for each** j **in** $\{j: (j,k) \in D, k = p\}$
    - $\alpha \leftarrow$ latest time predecessor j is currently scheduled to be completed
    - **if** $\alpha > \delta$
      - $R' \leftarrow R' \cup$ {set of resources currently allocated to work project j or a predecessor ancestor of j after time δ}
  - **while** $c_{\text{free}}(R', t) < c_p$
    - $r' \leftarrow$ resource in R and not in R' that has the earliest free time available for work allocation
    - $R' \leftarrow R' \cup \{r'\}$
  - **while** $c_{\text{total}}(p, R') < c_p$
    - $r' \leftarrow$ resource in R' that has the earliest free time available for work allocation
    - $\delta \leftarrow$ first time resource r' is free
    - **if** $t_{\text{predecessor finish}}(p) > \delta$
      - $j \leftarrow$ predecessor of p limiting project p to be started at δ
      - **if** $c_{\text{reallocate}}(j, r', \delta) \leq \varepsilon$ (ε representing a very small positive value)
        - Enlarge R' with any other resources that have work after δ in regard to project p's immediate predecessors or predecessor ancestors
        - Reallocate R' resource work after δ in predecessor constrained order with the full resource-group-context intensity through time
    - **else**
      - Assign resource r' to work project p up to time t or the amount of time-units the project still needs while only assigning work sequentially to time windows $W_{r'}$

FIGURE 2. SCHEDULING ALGORITHM PSEUDO CODE

3.4 Experiment Evaluation

In order to validate the effectiveness of the proposed method, the proposed method is compared to the variable-intensity task scheduling mixed integer linear programming (MILP) formulation that relies on discrete time steps for resource allocation, as proposed by (Askin 2003) and (J. Chen 2005). Li et al. (2015) employs a similar technique with the of use of additional continuous decision variables to denote the intensity of project work activity in a given discrete time period. The experiment considers multiple configurations of this MILP model by discretizing the planning horizon with a step size of either 1, 2, or 3 time units. The MILP method configurations are denoted as $MILP_1$, $MILP_2$, and $MILP_3$, the subscript indicating the discretization step size. The proposed methodology is denoted as *BILP* and discretize the planning horizon by the benefit deadlines. Every problem instance in the dataset, described below, is then solved and the total benefits for each approach and configuration are computed. IBM CPLEX 12.7 optimization software with the default settings is used to solve the problem instances formulated by both the proposed BILP models and the comparison MILP models. The optimizations are performed on a machine with an Intel Xeon 3.6 GHz processor and 32 GB of memory.

Both methods are applied to a dataset consisting of 1,800 problem instances. The dataset is based on the scheduling dataset provided by (Vanhoucke et al. 2008). Figure 3 shows an example problem instance formulation from the scheduling dataset's Patterson formatted files. Each problem instance holds 30 activities which are interpreted as optional projects. The successor activity relationship constraints are interpreted as designed by (Vanhoucke et al. 2008). The values in this dataset that indicate an activity's fixed duration are interpreted as the number of variable-intensity work units needed for project completion. The resources of the dataset are reinterpreted as events that contribute benefits with the completion of projects; each problem in the dataset holds four resources which are interpreted as events. The activity requirements for each of these resources are interpreted as the project's benefit measurement in regard to the associated event. Each event is provided a deadline in the order they were presented as 6, 12, 18, and 24 time units. The planning horizon is fixed at 24 time units.

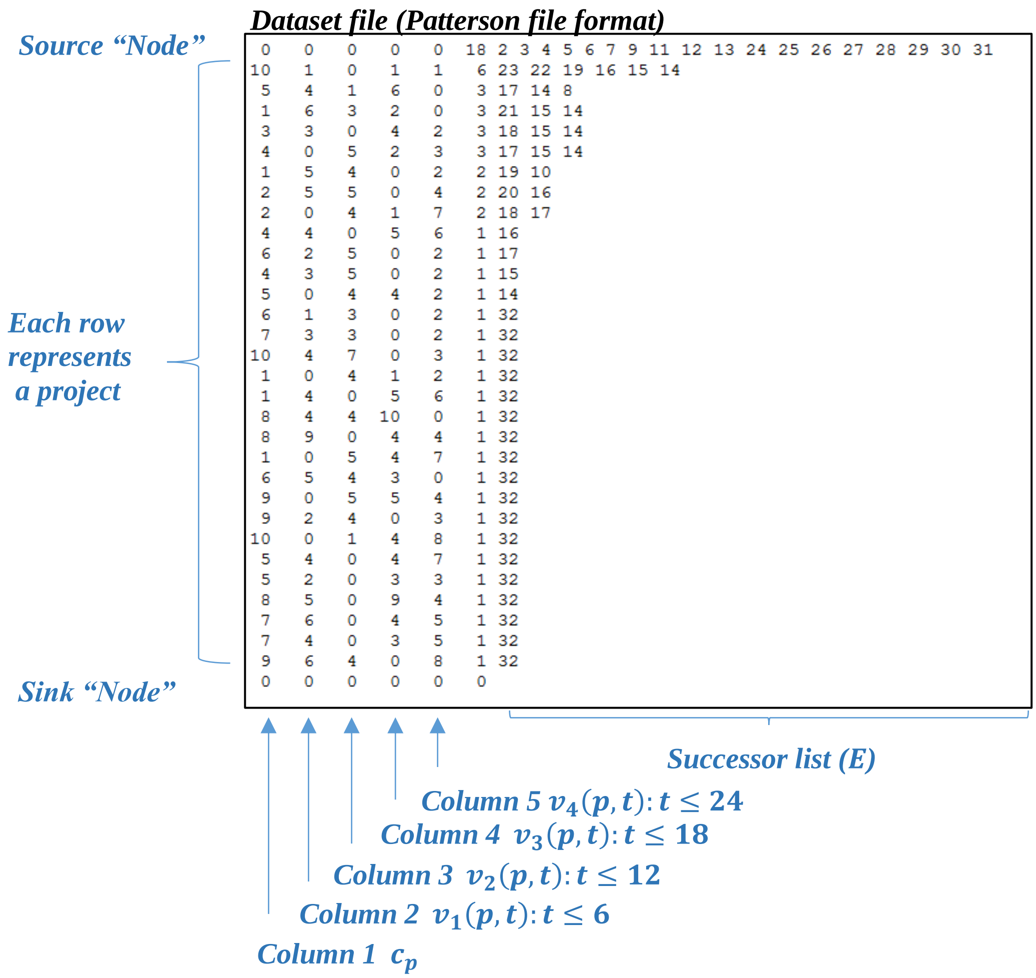


FIGURE 3. PROBLEM INSTANCE FORMULATION EXAMPLE

The proposed methodology's applicability to resource sizing analysis on a larger problem that consists of 300 projects is also demonstrated. For this demonstration, a randomly selected problem instance from the dataset generated by (Debels and Vanhoucke 2007) is used. This problem instance is interpreted as shown in Figure 3. The planning horizon is fixed at 24 time units. The demonstration considers three different resource profiles (i.e., resourcing alternative). The 'Status Quo' resource profile consists of 8 resources available over the whole planning horizon. The 'Part Time Support' resource profile adjusts this profile by adding 4 temporary resources for the first 12 time units. The 'Gradual Hire' resource profile adjusts the 'Status Quo' by adding (hiring) a single resource after each time period that would remain throughout the rest of the planning horizon.

3.5 Results and Discussion

The proposed methodology results in equal or better optimized values for all the problem instances in the benchmark dataset compared to all three configurations of the MILP-based method. Figure 4 shows the percentage of problem instances that the MILP method fails to discover a solution that matches the proposed method optimal solution's value. As the time-step size for the MILP model increases, the MILP methodology more frequently fails to find the *BILP* method's optimal solution. The average optimal value over the 1,800 problem instances for the proposed method resulted is 156, while the MILP average optimal values are 147 for $MILP_1$, 136 for $MILP_2$, and 125 for $MILP_3$. Notice the MILP time step size of 1 is less than or equal to any single project's work requirement in the dataset. An analyst is not guaranteed the problem's optimal solution if they use the MILP method with the smallest project work requirement as the MILP discretization step-size.

For the problem instances where the MILP methods fail to find an optimal solution, the distribution of the percentage of MILP methods' solution value compared to the proposed methodology's optimal value is depicted in Figure 5. As the time allocation step size of the MILP model increases, the optimal value differences increase.

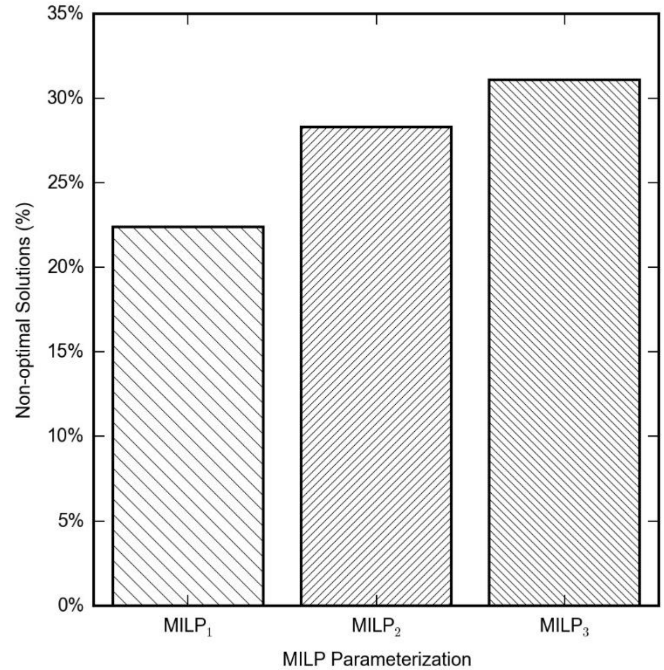


FIGURE 4. PERCENTAGE OF TIMES MILP METHOD FAILED TO FIND OPTIMAL SOLUTION

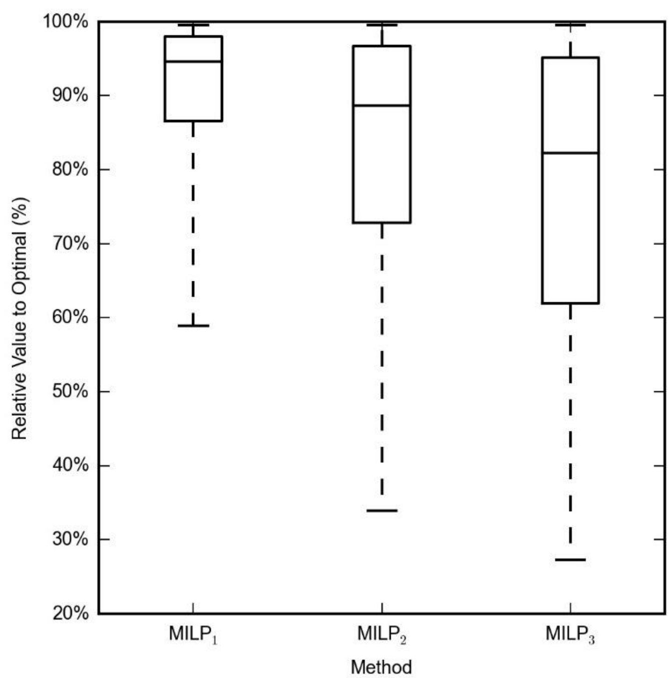


FIGURE 5. DISTRIBUTION OF SUB-OPTIMAL SOLUTION VALUES

The proposed methodology's solutions have more varied work intensity levels compared to the MILP solutions. The distribution of the schedule's average project work intensity (i.e., average number of resources assigned to a work activity through time) is shown in Figure 6 for the 1,800 problem instances. The average project schedule intensity for the proposed method is 2.5 resource entities, while the MILP average intensities are 2.86 resource entities ($MILP_1$), 2.27 resource entities ($MILP_2$), and 1.78 resource entities ($MILP_3$). In 66% of the problem instances, the proposed method creates a schedule with a lower average intensity than $MILP_1$.

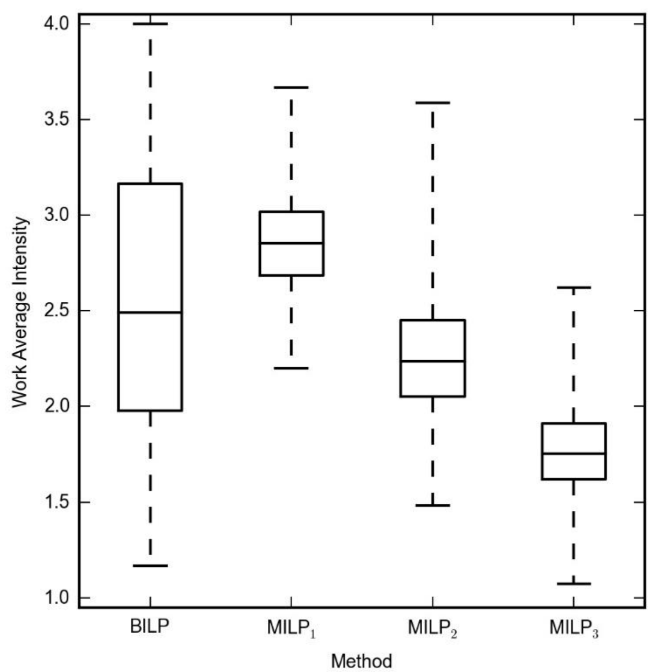


FIGURE 6. DISTRIBUTION OF WORK AVERAGE INTENSITY BY PROBLEM INSTANCE

The flexibility of the proposed methodology to increase the intensity of work to realize additional benefits is highlighted in problem instance 706 (one of the problems of which the BILP methodology generates a schedule using the maximum intensity possible). Figure 7 shows the resource to project assignment schedule for this instance using each methodology configuration. The projects are denoted ‘P-#’. The x-axis tick marks indicate the linear programming discretization of the time planning horizon. The problem instance’s successor/predecessor network is characterized by limited branching of a long project-to-project predecessor dependency sequence. The proposed methodology discovers and prescribes the maximum work intensity of 4 for every project selected. In other words, all four resource entities must work on the same project in time for every project in order to achieve the optimal solution while accounting for predecessor constraints. Notice that for the proposed BILP methodology, the start and end time of the work assignments mostly do not fall directly on the benefit deadlines. As the discretization step size increases for the MILP method, the solutions hold more gaps where the MILP method fails to find projects to assign to resources while meeting the projects’ predecessor requirements. The average resource utilization over all the problem instances are 98% for the proposed BILP method, and 94%, 88% and 82% for the MILP methods, respectively.

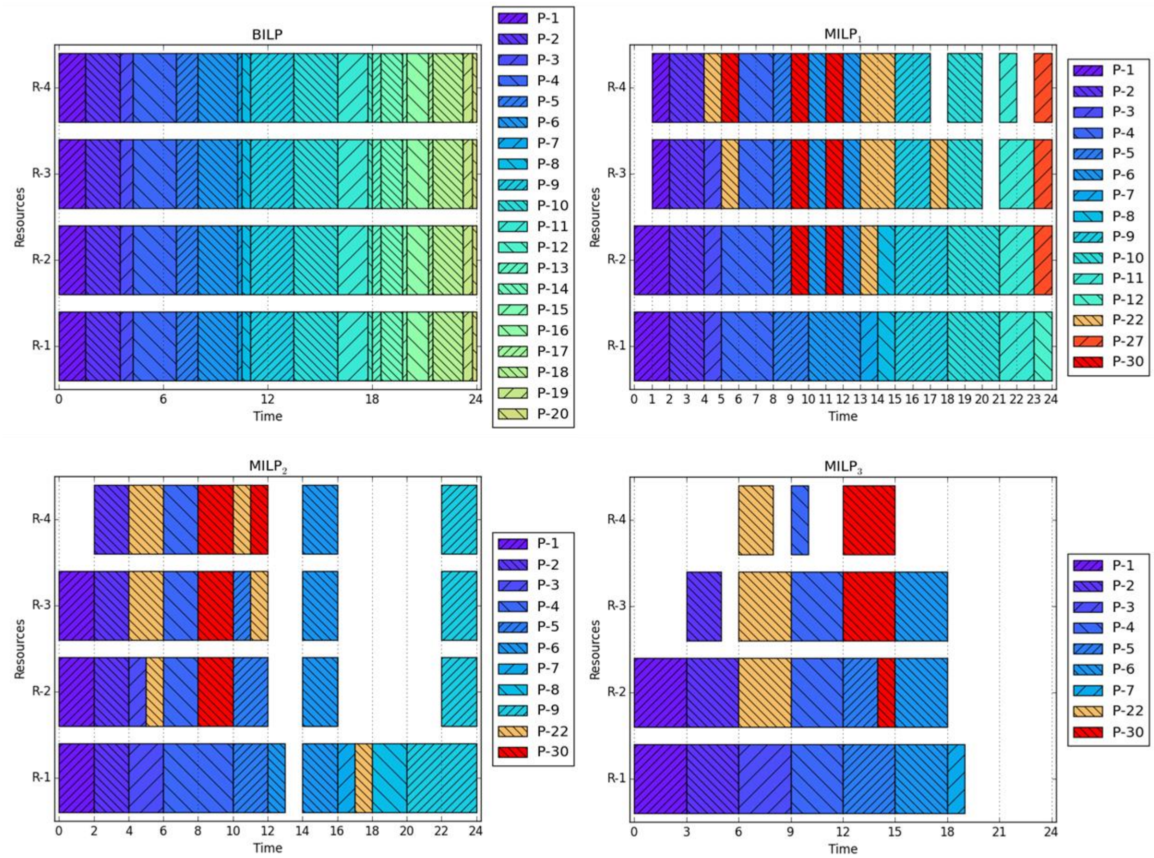


FIGURE 7. RESOURCE ASSIGNMENT SCHEDULE FOR PROBLEM INSTANCE #706 BY SOLUTION METHOD

The computation simplicity of the proposed BILP method results in a significantly quicker computation speed. The average computational speed given for the proposed methodology is 0.23 seconds, while the MILP average computational times are 5.29 seconds ($MILP_1$), 1.16 seconds ($MILP_2$), and 0.56 seconds ($MILP_3$). The proposed methodology on average finds an optimal solution in 11.3 seconds for the 300-project sized datasets, consisting of 480 problem instances from (Debels and Vanhoucke 2007); the $MILP_1$ method sometimes fails to resolve an optimal solution after days of computation for a single 300 project problem instance on the experimentation computer system.

To demonstrate cost-vs-benefit-vs-intensity analysis considering different resource profiles, a problem instance, #166, is used from the dataset (Debels and Vanhoucke 2007). Three resource profiles (i.e., options) are considered as described in section 3.4. The 'Part Time Support' option results in the most optimal realization of the project benefits, while the 'Status Quo' option results in the best benefit-to-resource ratio. The 'Gradual Hire' option results in a significant increase in work intensity to achieve its optimal solution. Figure 8 shows an overview of the optimal schedule solution characteristics by resource profile. Figure 9 shows a detailed view of the work schedule intensity. This type of intensity analysis adds insight for the organization to prepare the execution of the schedule in light of work intensity. Notice that almost 20% of the total scheduled work for the 'Gradual Hire' resource profile is assigned at an intensity level 12 (i.e., all 12 resource entities work on the same project at the given times). This highlights a limitation of the proposed BILP method compared to the MILP model as proposed by (Askin 2003; J. Chen 2005) in that the MILP method permits specified upper variable-intensity bounds. Solutions from the BILP method may require extra scrutiny to ensure the schedule meets work intensity efficiency assumptions, especially in regard to the consideration of adding resources.

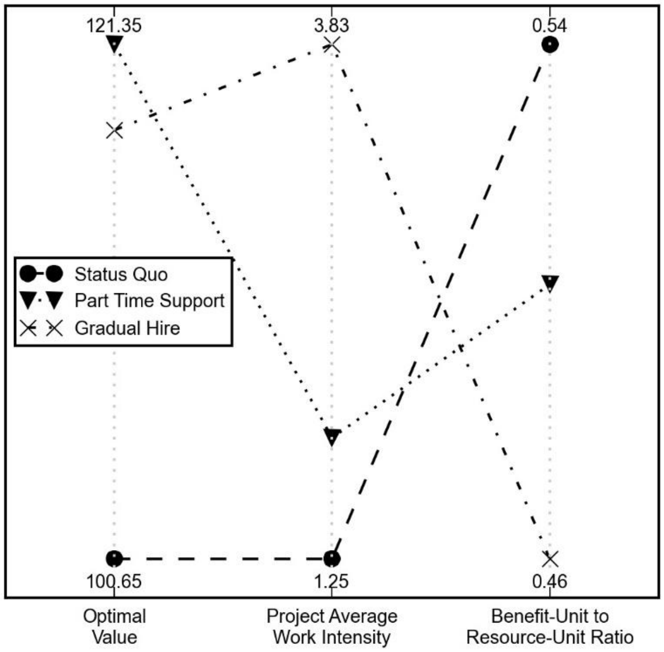


FIGURE 8. OVERVIEW OF BENEFITS AND INTENSITY BY RESOURCE PROFILE

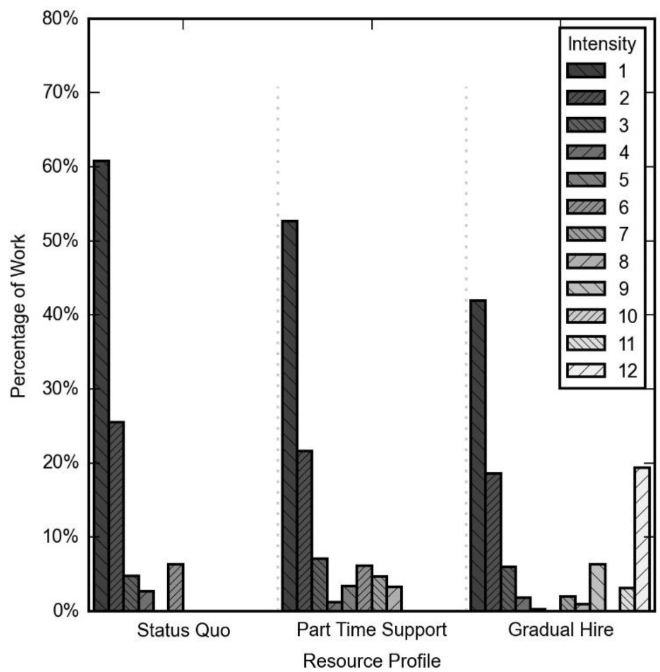


FIGURE 9. PERCENTAGE OF WORK AT INTENSITY LEVEL BY RESOURCE PROFILE

3.5.1 Additional Possible Extensions

The proposed methodology lends itself to extensions that can support additional problem characteristics applicable to other applications of the project selection and scheduling problem. The paper presents some of these possible extensions now.

To accommodate additional resource constraints, the methodology could incorporate one or many consumable resources into the first stage of the methodology by summing the amount of consumable resources available in each time period. This assumes that a known amount of consumable resources is promptly available at the start of each time period.

To address assignable resource unit efficiency differences through time, such as a new employee working slower than an experienced employee, the methodology could incorporate integrable functions of the assignable resources efficiency. The first stage of the methodology would compute the resources units available up to the end of the time period using an integration sum of the resource efficiency between the start of all time periods to the end of the given time period. The second stage's algorithm would require a few more adjustments to make properly time-consumption assignments and to ensure the calculations of expected accomplished work from project assignments matches the assignment's resources efficiency. Future research is recommended to investigate this concept further.

To address mandatory projects, constraint (3) could be modified to an equality constraint. To address no later than restrictions or no early restrictions, projects' decision variables could be removed from the model or restricted to zero for the appropriate time periods. To address project exclusivity restrictions, constraint set (7) could be incorporated to enforce the project mutual exclusion requirements where M_i denotes a set of mutually exclusive projects (i.e., at most one project $p \in M_i \subseteq P$ can be selected), indexed by $i = 1,2, \dots, m$.

$$\sum_{p \in M_i} \sum_{t \in T} x_{pt} \leq 1 , \forall i \in \{1, 2, \dots, m\} \tag{7}$$

To address project benefit interactions, placeholder, "dummy" projects could be incorporated to account for the benefit interaction effect. The dummy projects proposed would have predecessors of the projects that take part in the interaction effect and would have 0 work unit needs. To enforce the incorporation of negative interaction effects, the methodology would require the addition of activation enforcement constraints into the first stage binary integer linear programming model, such as constraint (8) that denotes interaction effect k (represented as a project with 0 work unit needs with predecessor edges to the projects holding the interaction) between projects i and j. Other model implementation variations have been proposed (Carazo et al. 2010; Stummer and Heidenberger 2003).

$$\sum_{t \leq t'} x_{it} + \sum_{t \leq t'} x_{jt} - 2 \sum_{t \leq t'} x_{kt} \leq 1, \forall t' \in T \tag{8}$$

3.6 Conclusions

Past project selection and scheduling optimization methods, while addressing project dependencies and variable intensity work, can potentially negatively constrain themselves from finding the optimal solution by disallowing predecessor projects to be completed in the same discrete time period and starting successor projects midway in same discrete time period. This paper proposes a project selection and scheduling methodology that removes the reliance of resource entity time allocations to the discretization of the planning horizon for a project selection and scheduling problem variant characterized by variable-intensity work, deadline defined benefits, and a resource pool of symmetrical skilled resources with availability changes in time. The paper demonstrates and compares the proposed methodology with a method that strongly relies on the discretization of the planning horizon for resource allocation. Based on a dataset of 1,800 problem instances composed of 30 projects, the results show the proposed methodology results in better solutions for over 20% of the problem instances compared to a MILP method configured for three different discretization step sizes.

The paper presents several ways to analyze the variable-intensity properties of a project selection schedule. The results show that even though the proposed methodology does not constrain the upper bound of work intensity, the methodology results in an average lower intensity for 66% of the problem instances compared to the unbounded variable-intensity method. The paper also demonstrates how the methodology can be leveraged to support resourcing decisions. The proposed BILP formulation takes less time to compute which allows more runs in time constrained environments to gain insight into sensitivity of resource inputs and the problem's objectives through interactive decision support systems.

Future research may be beneficial into the extension of the methodology to support unsymmetrical resource entity efficiencies, such as new employees that slowly improve work efficiency through the planning horizon. Research into the ability to incorporate the proposed methodology into methods that address more complex project selection and scheduling problems, such as multiple renewable resource types may also be beneficial. Given the relatively quick computational speed of the methodology, future research is suggested into extensions to the methodology to conduct more extensive resource allocation analysis and sensitivity analysis.

IV. Project selection considering re-production projects and aging products

4.1 Introduction

This paper presents a variant of the project selection and sequencing problem encountered by the Air Force, a methodology to address this problem, and a case study application of the methodology for an Air Force organization. In this variant, the product of the project is information. This paper defines information production planning (IPP) as the activity that prescribes what information to produce, when to produce it, and how to produce it (by what resources). The IPP problem of interest is to prescribe a production plan that maximizes the total value that information provides to multiple customers (i.e., requestors or users internal to an organization) with respect to resource constraints that limit fully satisfying every request. Organizations may confront the IPP problem in other decision situations including business marketing plans, investigative operations, and system development activities.

The problem possesses a few complicating attributes that limit the applicability of previously proposed methodologies. First, the problem includes product deterioration issues. Existing or to-be-produced information products may become stale by the time of usage. Given a multi-year planning horizon, the planning process also includes production decisions on to whether to refresh (i.e. reproduce) the information product. Second, the information products being considered for production can inform multiple customer decisions (not to be confused with the planning production decisions). These usage-decisions may have different time use profiles and different objectives. For example, a customer may want information for a system engineering test-event being performed next year, another customer may want the same information over the course of the next three years to support ongoing system reprogramming decisions, while another customer may want the information immediately for early system requirement definition.

This paper proposes a methodology to prescribe production plans consisting of the selection and the scheduling of discrete information production projects over a planning horizon given stakeholder defined budget alternatives. This paper first presents a literature review, followed by the methodology, a case study employing the methodology, the results of the case study, and conclusions.

4.2 Literature Review

The IPP problem of interest faces product deterioration issues, reproduction decisions, and products that support non-monetary objectives. Recent efforts have studied information production planning from the perspective of crowdsourcing (Basu Roy et al. 2015; Bessai and Charoy 2017; Mavridis, Gross-Amblard, and Miklós 2016; Rahman et al. 2015; Sinha, Majumder, and Manjunath 2016; Tran-Thanh et al. 2014). For example, Roy et al. present an integer linear programming approach to optimize information production planning given crowdsourcing uncertainties and task attributes (Basu Roy et al. 2015). A limitation of the discovered crowdsourcing research is that it does not address optimization in light of information deterioration issues.

Research exists concerning the deterioration and perishability of products in material based production problems (Amorim, Günther, and Almada-Lobo 2012; Pahl and Voß 2014; Rong, Akkerman, and Grunow 2011). The characteristics of the IPP problem's deterioration, namely a shareable-digital product and different deterioration criteria for different customers' product usage, are not addressed by the discovered supply chain production planning research. The proposed methodology builds upon the value (i.e., benefit) measurement method introduced by (Golabi, Kirkwood, and Sicherman 1981) to account for temporal objective effects due to information product deterioration and different customer usage timelines. The methodology integrates a mathematical modeling formulation similar to techniques proposed in R&D project selection research (Carazo et al. 2010; Stummer and Heidenberger 2003) to support the optimization of when selected "projects" are produced in regard to time and integrates an extension to address reproduction decisions in the mathematical model.

4.3 Methodology

The methodology is a 7-step process. Figure 10 shows an overview of the process. The first step of the methodology discretizes the planning horizon to adequately account for resource levels through time and for the temporal benefit effects from production decisions. For example, the case study discretizes the planning horizon based on the organization's fiscal year which provides a clear link to resource budgets, while also providing support to temporal planning issues that address deterioration of information products. The second step of the methodology is to gather the information requests and to estimate the value of information (VOI) for each information request's usage decision. Let R represent the set of information production requests and r represent an element of this set. Let X represent the set of requested information products and x represent an element of the set. Let R_x denote the set of information production requests for product x. Let there be a many-to-one relationship between requests to information products (i.e., a request specifies a single product, multiple requests may specify the same product). The VOI represents the expected value change in outcomes for the requesting decision context if the information is procured and used compared to if the information is not procured and used. Decision trees are one way to compute the VOI (Lawrence 2012). Unlike the decision tree VOI approach, this step of the proposed methodology defers the information acquisition decision and cost considerations to a later step to support the holistic cost-vs-benefit analysis considering potentially multiple users of the same information and other uses of limited resources. During this step, the requests' use of information over time are considered. For continuous information-usage decision contexts, these VOI measurements are considered from the perspective of an average point in time over the length of the mathematical model's discrete time periods and for each time period if the VOI changes with respect to time. Let $v_{r,x}(t)$ represent request r's VOI for having the information product x at time t.

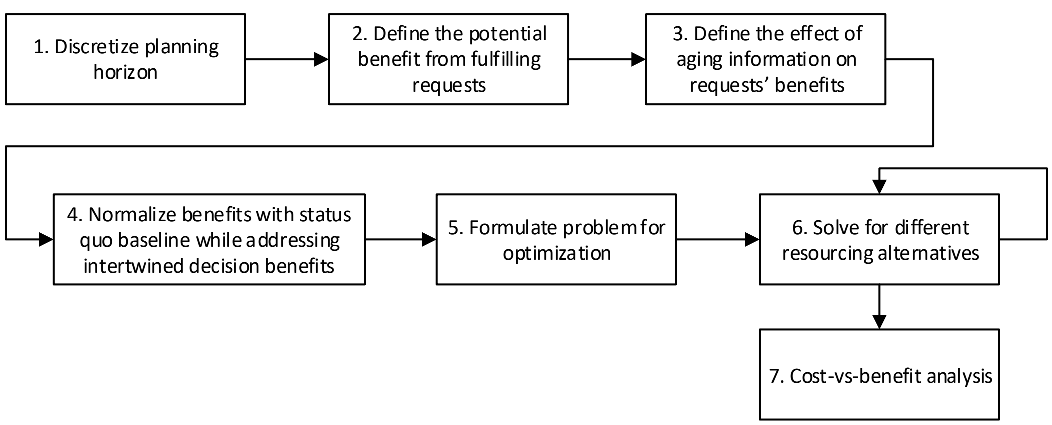


FIGURE 10. INFORMATION PRODUCTION PLANNING METHODOLOGY FLOW CHART

The third step is to model the effect of information aging (i.e., deterioration) on each request's VOI. This step can be executed with the first step, incorporating discrete decision tree scenarios to represent different ages of information. Taking the discrete values tied to discrete ages, a deterioration factor function could be fit. This step also can be executed as a separate study of historical trends regarding similar decision contexts and information aging states. An approximation method based on a subject matter expert (SME) estimates is also possible. For example, the SME method may elicit the likely age of information when the information begins to inaccurately inform the decision context and the likely age of information that the information is no longer of value for the decision. Using these two inputs, a linear approximation of the degrading VOI may be constructed over the possible information product age range. This approximation technique assumes the change of VOI through time is independent of changes at static points in time. The methodology can support expected, explicit changes at static points in time by anchoring age of information value estimation functions to an absolute time and date, though the case study does not demonstrate this capability. Let $d_{r,x}(a,t)$ denote the deterioration factor function (i.e., the percentage of request r's value that is expected to be realized due to information product x being of age a at time t). Figure 11 shows two examples of how deterioration may affect the VOI in regard to a and independently of t.

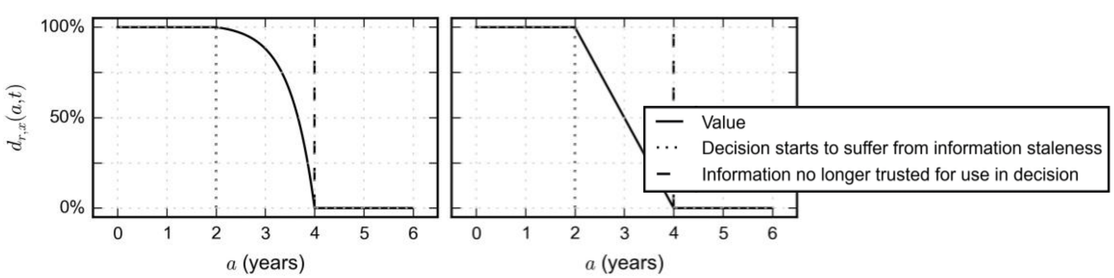


FIGURE 11: EXAMPLES OF A REQUEST'S VALUE CHANGES AS THE REQUESTED INFORMATION PRODUCT AGES

The forth step is to normalize the decisions' VOI measurements to addable values. Note that the units of VOI measurement may change between requestors' decision contexts, especially in regard to non-profit oriented organizations or requirement centric system engineering environments. A common measurement unit used in profit-oriented organizations is net present value (NPV). To address non-profit oriented objectives, the normalization process utilizes (and validate the assumptions of) measurable value functions as necessary (Golabi, Kirkwood, and Sicherman 1981; Liesiö 2014). Let $v_m(v_{r,x}(t))$ denote the normalized value for having the information product *x* at time *t* with respect to request r. In addition, either approach in the problem of interest requires the consideration of a baseline (Liesiö and Punkka 2014). For example, consider an information product that was recently produced and a decision maker evaluating whether to produce it again over the planning horizon. Reproductions of this information product early within the planning horizon cut short the value of past productions since the potential future value of the information would already be realized without the re-production of the product early in the planning horizon. The value of past productions, contributing to the value function baseline, must be subtracted from potential future re-productions aligned in time to accurately capture just the additional benefits of the production event(s). Figure 12 shows how the valuation of information production incorporates this adjustment with the deterioration factor for a continuous information request assuming $v_m(v_{r,x}(t))$ normalizes the benefit of fulfilling a request to a zero-to-one scale. Without subtracting the benefit realized by the past productions, information-product reproduction options would inaccurately include the benefits already realized from past productions. The potential result is the selection of reproduction options that result in little or no additional value.

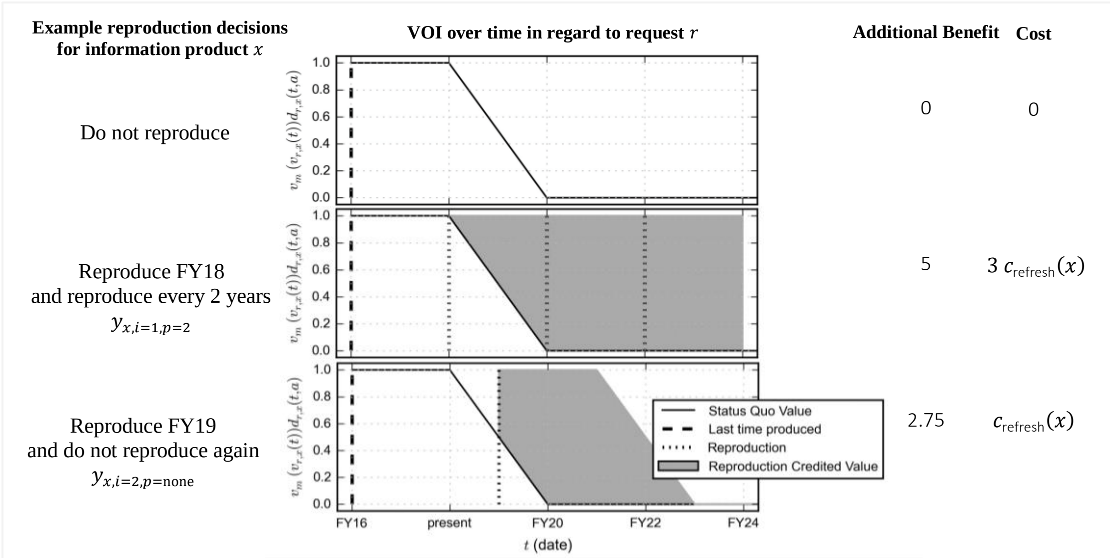


FIGURE 12: SUBTRACTING STATUS QUO EXPECTED BENEFITS TO COMPUTE THE BENEFIT OF REPRODUCTION(S) FOR A CONTINUOUS INFORMATION USAGE REQUEST

With these inputs, the IPP problem is to choose when to produce or reproduce the requested information products over the planning horizon to maximize the total value of information, as determined by the sum of the requests' deterioration-adjusted VOI, given resource constraints that limit production activities to fully satisfy all the information requests. The fifth step mathematically models this IPP problem. Similar to past project selection and scheduling approaches facing a continuous time decision space, this paper introduces an approximate discrete optimization formulation that discretizes the continuous, time planning horizon while accounting for the value of information time interaction effects at these discrete times.

Let T denote the set of end times for each time period dictated by the discretization strategy. Let g_x denote the age of information product x at the start of the planning horizon. To simplify the formulation without loss of representativeness, information products that have not been produced in the past are given a very large g_x

value to reflect the current absence of the information product. Let $y_{x,i,p}$ denote the binary decision variable to first produce information product x by time i, $i \in T_x \subset T$, and refresh (reproduce) the product every p time periods, $p \in P_x$, to the end of the planning horizon, where T_x denotes the set of possible first completion times being considered for product x and where P_x denotes the set of refresh policies (number of time periods between re-productions) being considered for product x. Let $U_{x,i,p}$ represent a set of tuples spanning the planning horizon based on production events as specified by i and p, where the tuples (i.e., elements) of $U_{x,i,p}$ take the form of (j, k, a) such that j denotes the time a production event is to be completed or the start of the planning horizon, k denotes the next time a reproduction event occurs or the end of the planning horizon, and a denotes the age of information product at the start of the time interval. Note that a holds the value of zero for each tuple except for the tuple representing the first time period at which a holds the value of g_x. Let $b_{x,i,p}$ denote the normalized sum of additional benefits realized for making this decision as denoted in equation (9). Note that the status quo value is subtracted in this equation to derive just the additional benefits from re-productions of an information product.

$$b_{x,i,p} = \sum_{r \in R_x} \sum_{(j,k,a) \in U_{x,i,p}} \int_j^k v_m\big(v_{r,x}(t)\big)\, d_{r,x}(a + (t - j), t) - v_m\big(v_{r,x}(t)\big)\, d_{r,x}(t - g_x, t)\, dt \qquad (9)$$

Like the benefits, the resource costs are calculated in regard to time. Let $c_{\text{first}}(x)$ denote the estimate cost to (re)produce the information product x for the first time in the production plan. Let $c_{\text{refresh}}(x)$ denote the estimate cost to reproduce the information product x, also known as the refresh cost. Let $c_{x,i,p,t}$, as defined by equation (10), denote the estimated cost levied onto time period t if $y_{x,i,p}$ is selected. Let r_u denote the sum of the resource units available up to time u.

$$c_{x,i,p,t} = \begin{cases} c_{\text{first}}(x) & i = t \\ c_{\text{refresh}}(x) & (t-i) \bmod p = 0 \text{ and } (t-i) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

With these definitions, the following binary integer linear programming model is developed. Let equation (11) denote the objective function. Let equation (12) denote the resource constraints for each time period. The resource constraints assume unused resources from time periods can be utilized for later time periods. The resource constraints can be replicated to support additional consumable resource types. Let equation (13) represent the mutual exclusive policy and production options for each information product. Let equation (14) represent the binary constraints of the decision variables.

$$\max \sum_{x \in X} \sum_{p \in P_x} \sum_{i \in T_x} b_{x,i,p} y_{x,i,p} \tag{11}$$

subject to

$$\sum_{x \in X} \sum_{p \in P_x} \sum_{i \in T_x} \sum_{t \in \{t : t \in T,\ t \leq u\}} c_{x,i,p,t}\, y_{x,i,p} \leq r_u \text{ , } \forall u \in T \tag{12}$$

$$\sum_{p \in P_x} \sum_{i \in T_x} y_{x,i,p} \leq 1 \text{ , } \forall x \in X \tag{13}$$

$$y_{x,i,p} \in \{0,1\}, \forall p \in P_x, \forall i \in T_x, \forall x \in X \tag{14}$$

Finally, the sixth and seventh step of the methodology solves and presents the mathematical model's optimal solutions (representing production plans) for different budgeting alternatives explicitly defined by the stakeholders to a decision maker.

4.4 Case Study

The case study consists of eleven information customers within the Air Force making in total 545 requests for 407 information products for use over the next 6 years for system engineering and system reprogramming purposes. Production management cost models estimate an up-front production cost and a refresh cost for each requested information product. The year time-unit discretizes the planning horizon aligning to the organization's fiscal year. The stakeholders incorporate a VOI scale, composed of criticality levels (CL), and assigns each request a CL to represent the VOI for the request. A "CL 1" request represents the highest priority need. A "CL 4" request represents the lowest priority need. The $v_{r,x}(t)$ value function returns the request's criticality level value if product *x* is wanted at time *t* and 0 otherwise. With inputs from subject matter experts and organization leaders, the v_m value function normalizes the values from the CL scale according to the tradeoff preferences. Specifically, the "CL 4" requests are mapped to a value of 1, the "CL 3" requests are mapped to a value of 2 (two "CL 4" requests are equally preferred to one "CL 3" request), the "CL 2" requests are mapped to a value of 5, and the "CL 1" requests are mapped to a value of 20. 168 of the requests are single point-in-time information usage requests (use the information in only a single time period). The remainder of the information requests are continuous requests and include a desired refresh rate varying from one year to five years. Deterioration factor functions model the effect of information-product-aging on value. For the continuous usage requests, the full value of the request is realized for each time period as long as the product's age is less than the request specified refresh rate. Otherwise, partial value of the request is realized according to a linear degradation through time until it is considered completely out-of-date (if a product is older than twice the desired refresh rate). For the single point-in-time usage requests, if the product already exists or will be created before the request specified need (i.e., usage) date then the request realizes the full VOI amount. Otherwise, no value is realized. Two product refresh (i.e., reproduction) policies are considered: 1) the minimal refresh rate of the requests concerning a product; and 2) none (only one production over the planning horizon or only one reproduction if already produced).

Six resource configurations (i.e., budget alternatives) are considered. The budget alternative labeled "100% (baseline)" denotes the minimum yearly budget needed to fully satisfy all the requests (approximately $14 million dollars over the 6 year planning horizon). The other five alternative budget configurations consider a percentage of this baseline budget. For each budget alternative, CPLEX 12.7 software (with the default settings) on an Intel 3.6GHz processer computer and 32 GB of memory solves the problem models and generates a production plan for each budget alternative.

4.5 Results

CPLEX finds the mathematical model optimal production plans for the six considered budget alternatives in 50 seconds. Figure 13 show a cost-vs-benefit plot of these production plans where the x-axis represents the cost (i.e., total budget sum over the 6 year planning horizon) and the y-axis represents the benefit (i.e., percentage of the information requests total value realized by the production plan). The 33%-of-baseline budget alternative's production plan realizes approximately 90% of the information requests' value. The gradual change in percentage of information value realization going from the baseline budget down to the 33%-of-baseline budget suggests several relatively less valued, but very costly information products were requested.

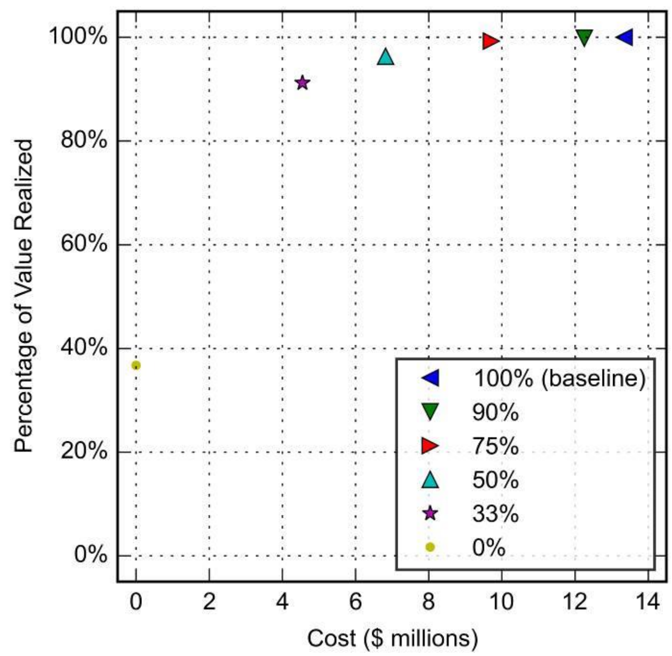


FIGURE 13: PRODUCTION PLANS COST-VS-BENEFIT PLOT

Figure 14 shows the percentage of total value realized over time given the different production plans. If no additional production takes place from the present time, the results suggests approximately 10% of the information requests' value is realized 6 years from FY17 due to information becoming stale and the lack of additional production or reproduction of requested information products. Figure 15 shows value gaps from the perspective of the case study's VOI criticality levels (CL) for the 33% and 50%-of-baseline alternative's optimal production plan. Each row in this figure represents a requested information product. The color of each cell represents the size of the value gap caused from the production plan given the budget and all of the information requests relevant at any given time. The dark blue color, not applicable ("NA"), represents the time a product is not expected to be needed by any request. The 33%-of-baseline production plan results in few CL 1 gaps while mostly satisfying the majority of the information product needs through time. Furthermore, the 50%-of-baseline production plan prevents the occurrence of CL 1 value gaps. As the budget increases from 50%, only CL 2 or smaller gaps remain in the optimal production plans given the budget constraints.

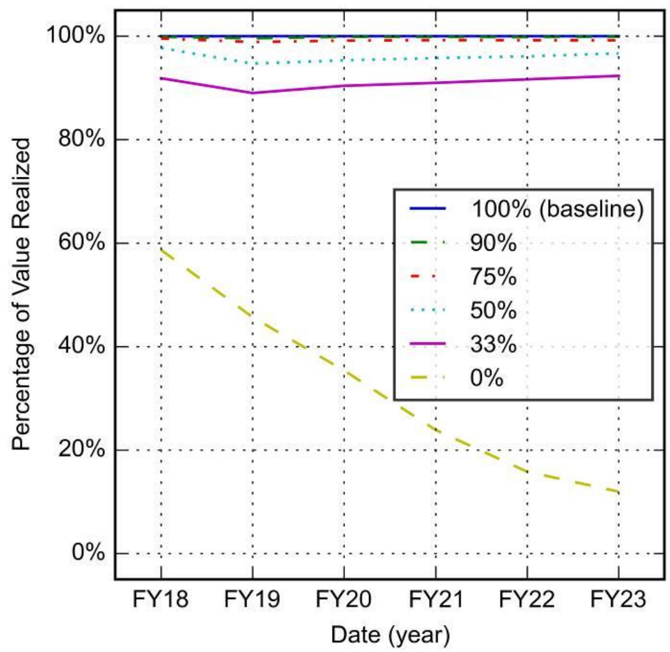


FIGURE 14. BENEFIT THROUGH TIME BY PRODUCTION PLAN PLOT

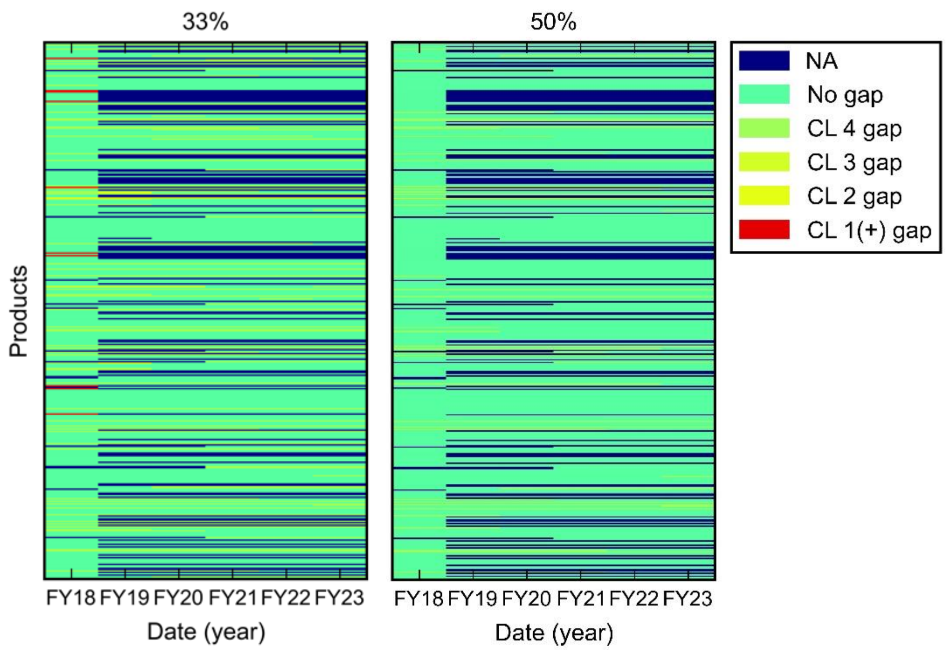


FIGURE 15. VALUE GAPS THROUGH TIME PLOTS

4.6 Conclusions

This paper introduces an information production planning problem with product-value-deterioration attributes and multiple customers holding temporal objectives. This paper then proposes a methodology to support the creation of a production plan that maximizes the value the produced information contributes to the organization's objectives while addressing: 1) how to evaluate information production decisions and reproduction policies considering the benefit to multiple information customers' requests in time; 2) how to incorporate the deterioration of information products into an production benefit evaluation method; and 3) how to mathematically model the information production decisions to support production planning optimization. A case study incorporates the methodology and shows the methodology's computation tractability for an instance of this problem encountered at a defense organization. The study shows numerous cost-vs-benefit insights including approximately $5 million dollars of the $14 million dollar fully funded budget realizes over 90% of the requested information potential value and 50% of the fully funded budget production plan prevents the occurrence of the most critical value gaps. Future research is suggested into semi-automated methods to perform sensitivity analysis on the funding levels and into methods to support ongoing planning processes incorporating past production plans.

V. Addressing modeling inadequacies

Elusive-to-model preferences and problem attributes hinder decision makers directly accepting a mathematical model's optimal solution for project portfolio and similar resource allocation problems with many near-optimum alternatives. Past research proposes generating a small, decision-space-diverse, set of near-optimum alternatives and letting decision makers leverage their elusive-to-model knowledge to choose or to base an alternative from this set through a decision support system. This paper highlights a gap in past research to address a correlation attribute of this decision-space-diversity objective for portfolio selection problems. The paper proposes an extension to a prevalent alternative generation technique to address this correlation attribute of diversity and applies the technique to a fantasy sports portfolio problem's binary linear programming optimization formulation. An extensive numerical experiment is performed for multiple technique settings and problem sizes with the problem instances developed from parameter distributions displayed in a real-world dataset. The results show the proposed extension significantly increases diversity when correlations of the generated solutions' attributes are considered.

5.1 Introduction

The portfolio problem consists of the selection of a subset of conceivable elements (e.g., projects) to maximize objectives with limited resources and technical constraints (Kleinmuntz 2007). For a portfolio selection problem with β number of selectable elements (e.g., projects), a decision maker must choose one portfolio (i.e., a combination of projects) from 2^{β} possible projects combinations, ignoring technical and resource constraints that restrict many of these alternatives. Mathematical optimization models are commonly employed to support portfolio decision problems (A. Salo, Keisler, and Morton 2011).

Even with rigorous mathematical optimization modeling, decision makers often do not immediately choose the first, mathematically-model optimum solution given elusive-to-model problem considerations (Hennen et al. 2017; Voll et al. 2015). Approaches to handle modeling or information inadequacies include interactive decision support systems (Argyris, Figueira, and Morton 2011; Nowak 2013; da Silva et al. 2017) and preference programming (Fliedner and Liesiö 2016; Liesiö, Mild, and Salo 2007; Pape 2017). An approach suggested by E Downey Brill, Chang, & Hopkins (1982) is expanded upon in this paper. E Downey Brill et al. suggests the presentation of multiple, decision-space-diverse near-optimal solutions to a decision maker to address these elusive-to-model problem considerations. The reasoning is that the explicit differences manifested in the suggested solutions provides a tangible bridge or spark for a decision maker to employ their elusive-to-model knowledge. E Downey Brill et al. labels the process of alternative generation "Modeling to Generate Alternatives" (MGA) since the process employs additional modeling to discover decision-space-diverse solutions among the potentially large set of near-optimal solutions. This paper denotes a MGA technique's result (e.g., a set of near-optimal portfolios to present to a decision maker) as a solution set. Figure 16 depicts the relationships of these components used within a decision support system (DSS) for a portfolio decision problem.

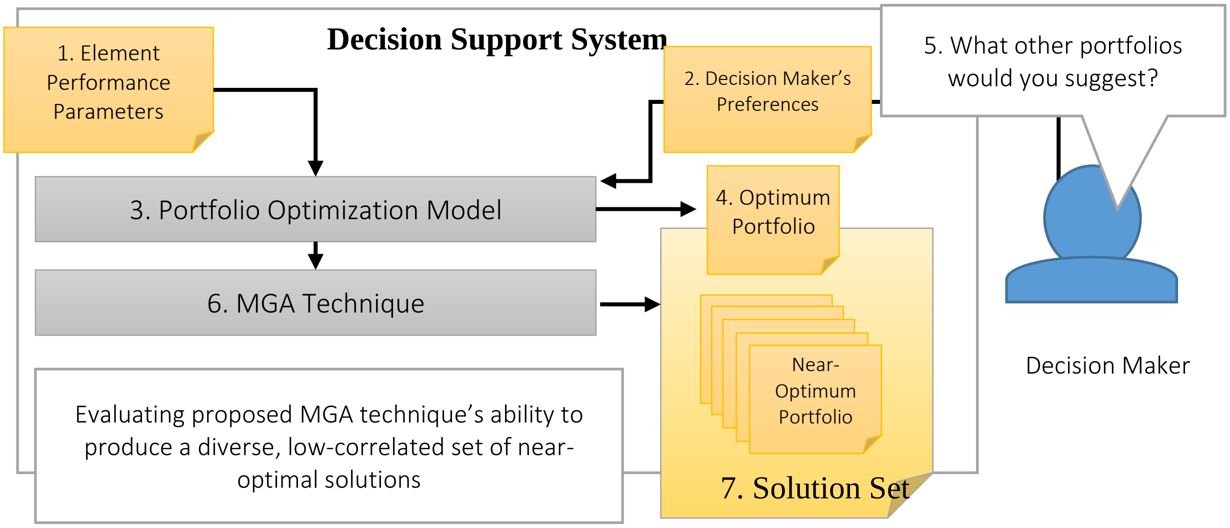


FIGURE 16. OVERVIEW OF COMPONENTS AND CONTEXT OF EFFORT

The objectives of MGA techniques are to produce diverse solution sets (e.g., more diversity between the suggested solutions leads to more opportunities for a decision maker to recognize and ultimately address any un-modeled preferences or other technical interactions between the decision variables in a solution) and that these solutions in this set are nearly-optimal from the perspective of the original optimization model's objective function. This paper proposes and evaluates an extension to the flexible Hop Skip Jump (HSJ) MGA technique (E Downey Brill, Chang, and Hopkins 1982; DeCarolis et al. 2016) to address a correlation attribute of this diversity objective in portfolio problems. HSJ takes an original optimization model, equations (15), and initial solution (such as the optimal solution to the original optimization model) and solves an adjusted model, such as model (16), where X represents the feasible decision space, x represents a feasible multi-variate decision, $f(x)$ represents the original objective function. Next, an adjusted model is iteratively solved to generate additional solutions different from the previously generated solutions. Model (16) represents a simple form of this adjusted model where K

denotes the set of non-zero decision variables in past solutions (including the initial solution) and T represents a near-optimality threshold. The adjusted model is solved iteratively until stopping criteria is met such as generating a desired number of solutions. The minimization of past, non-zero decision variables causes decision-space-diverse solutions while the $f(x) \geq T$ constraint enforces near-optimality. Brill et al. allude to additional forms of the adjusted model's objective function including weighting the decisions variables. For example, DeCarolis, Babaee, Li, & Kanungo (2016) study two adjusted HSJ adjusted model formulations employing decision variable weights with respect to an energy system optimization problem.

$$\max f(x), \quad \text{subject to } \; x \in X \tag{15}$$

$$\min \sum_{k \in K} x_k \quad \text{subject to } \; f(x) \geq T, x \in X \tag{16}$$

Note that HSJ techniques provide desirable features for use in portfolio decision support systems. HSJ techniques can be applied to any decision support system modified linear programming optimization formulation without extensive calibration, such as the binary linear programming formulation of the portfolio selection problem presented in Section 3. HSJ techniques only require a threshold for near-optimality input to extend the original optimization model. Also, decision support systems (DSS) utilizing HSJ can generate and present additional solutions in a piecemeal fashion as desired by the DSS user.

The HSJ technique can be classified as a greedy heuristic approach to generate decision-space-diverse solution sets. The original HSJ technique executes iteratively and at each iteration generates a near-optimal solution that minimizes the selection of past elements without considering how the elements are combined (i.e., selected together) in past solutions; this potentially can cause unnecessary element selection combinations (i.e., increased selection correlations) in the solution set for portfolio based problems. The authors qualitatively observed this result in testing datasets and this paper quantitatively evaluates this limitation compared to a HSJ extension proposed in Section 5.2. Related MGA research, (Baugh Jr., Caldwell, and Brill Jr. 1997; S.-Y. Chang, Brill, and Hopkins 1982; Loughlin et al. 2001; Shir et al. 2010; Zechman and Ranjithan 2007), does not explicitly consider this aspect of diversity through the use of measures not affected by solution set correlation such as the pair-wise average distance. Equation (17) defines pair-wise average distance, denoted ${}^{\mathrm{PWA}}D$, where $S\ is\ an\ n\ x\ q$ matrix and $S_{k,i} = \begin{cases} 1\ if\ element\ i\ is\ in\ portfolio\ k \\ 0\ otherwise \end{cases}$.

$$ {}^{\mathrm{PWA}}D(S) = \frac{\sum_{i=1}^{q}((\sum_{k=1}^{n} S_{k,i})(\sum_{k=1}^{n} 1 - S_{k,i}))}{\binom{n}{2}}, \qquad (17) $$

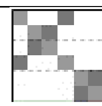
To demonstrate this issue, consider two solution sets generated from MGA techniques. Both sets contain four solutions. The first solution set contains two portfolios with elements A & B and two other portfolios with elements C & D. The second solution set contains one portfolio with elements A & B, another portfolio with elements C & D, one portfolio with elements A & C, and one portfolio with elements B & D. From a pair-wise average distance diversity perspective, these solution sets are equally diverse. Incorporating correlation attributes as part of the diversity criteria, the second set is more diverse and is more preferred given the smaller covariance. The off-diagonals elements of the covariance matrix, equation (18), represent the selection covariance for solution sets, where $s_{i,j}^2$ is calculated using equation (19). Table 3 presents the diversity and covariance measures of these solution sets (i.e., the more covariance in the covariance heat map demonstrates the increased correlation of the first solution set). The columns of S represent the elements A, B, C, and D, followed by two other elements.

$$\mathrm{cov}(S_{n\times q}) = \begin{bmatrix} s_{1,1}^2 & s_{1,2}^2 & \cdots & s_{1,q}^2 \\ s_{2,1}^2 & s_{2,2}^2 & \cdots & s_{2,q}^2 \\ \vdots & \vdots & \ddots & \vdots \\ s_{q,1}^2 & s_{q,2}^2 & \cdots & s_{q,q}^2 \end{bmatrix} \tag{18}$$

$$s_{i,j}^2 = \frac{1}{n-1}\sum_{k=1}^{n}(S_{k,i} - \mu_i)(S_{k,j} - \mu_j) \tag{19}$$

where μ_i is the average value of the i^{th} column of S.

TABLE 3. SOLUTION SET COVARIANCE MEASUREMENTS WITH RESPECT TO PAIRWISE AVERAGE DISTANCE DIVERSITY MEASURE

S	${}^{\mathrm{PWA}}D(S)$	Covariance Matrix Heat Map	${}^{\mathrm{SP}}_{\theta=0.5}D(S)$	${}^{\mathrm{I}}D(S)$
[1,1,0,0,1,0], [1,1,0,0,0,1], [0,0,1,1,1,0], [0,0,1,1,0,1]	4		2.58	1.38
[1,1,0,0,0,1], [1,0,1,0,1,0], [0,1,0,1,1,0], [0,0,1,1,0,1]	4		2.84	2.38

Ulrich et al. (2010) highlights similar limitations of the pairwise average difference measurement approach to quantify diversity of MGA techniques' results based on continuous decision variables and suggests the use of the Solow & Polasky diversity measure (Solow and Polasky 1994). The Solow-Polasky measure, denoted ${}^{\mathrm{SP}}D(S)$, recognizes increased correlations in portfolio solution sets and successfully distinguishes the increased diversity of the second solution set as demonstrated in Table 3. See Section 5.4 for the formulation of this correlation sensitive diversity measure. This papers uses both the ${}^{\mathrm{SP}}D(S)$ measure and a proposed information entropy based measure, ${}^{\mathrm{I}}D(S)$, in the evaluation of the diversity objective of interest.

The paper proceeds as follows: Section 5.2 formally presents the portfolio problem and optimization model, the original HSJ technique tailored to this portfolio problem, and a proposed HSJ extension technique. Section 5.3 introduces an application of the portfolio optimization problem and HSJ techniques in regard to a fantasy sports decision support system. Section 5.3 also includes a discussion of the complexity of the problem given the application problem. Section 5.4 provides the details for the numerical experiment and the correlation sensitive diversity measures. Section 5.5 presents the results and a discussion of the experiment's outputs.

5.2 Methods, HSJ and proposed extension

This paper studies HSJ techniques with respect to a parametrized portfolio problem (F. Li et al. 2012). In this problem, the decision maker must select a subset of discrete elements (binary yes/no decisions) to maximize the benefits the elements produce given deterministic resource constraints and element benefit estimates. The problem is modeled below with a binary linear programming optimization formulation where

Table 4 presents the notation used in this formulation, equation (20) denotes the objective, equation (21) denotes the selection constraints given a budget, equation (22) denotes the number of elements that must be included in any portfolio solution, equations (23) represents a set of technical constraints on the parameters of each project, and equation (24) denotes the binary nature of the decision variables.

$$\max \sum_{i=1}^{q} b_i x_i \tag{20}$$

Subject to:

$$\sum_{i=1}^{q} c_i x_i \leq C \tag{21}$$

$$\sum_{i=1}^{q} x_i = m \tag{22}$$

$$\sum_{i \in I_{p=v}} x_i \leq (\geq)\, l\,, \forall (p, v, l, o) \in L \tag{23}$$

$$x_i \in \{0, 1\} \tag{24}$$

TABLE 4. OPTIMIZATION MODEL NOTATION

Symbol	Description
I	set of selectable elements described by a δ-tuple, indexed by $i = 1, \dots, q$; $I = \{(1, \dots, \delta)_i \mid i = 1, \dots, q\}$ and where i_p is the p^{th} parameter of the i^{th} tuple
$I_{p=v}$	set of indices i such that the p^{th} parameter of the i^{th} δ-tuple is equal to v; $I_{p=v} = \{i \mid (1, \dots, \delta)_i \in I \text{ and } i_p = v\}$
x_i	binary decision variable for selecting i^{th} tuple of I
b_i	estimated benefits of selecting i^{th} tuple of I
c_i	cost of selecting i^{th} tuple of I
C	budget
m	number of elements that the portfolio must contain
L	set of 4-tuples (p, v, l, o) that define values of interest v for the p^{th} parameter of $(1, \dots, \delta)_i \in I$; $L = \{(p, v, l, o)_j \mid j = 1, \dots, r\}$, where l is a bound on the number of elements $(1, \dots, \delta)_i \in I$ that can be selected with this parameter value and o is the direction of that bound, $o = 0$ denotes $\leq$ and $o = 1$ denotes $\geq$
V_p	set of unique p^{th} parameter values of tuples in I; $V_p = \{\text{distinct } i_p \mid i = 1, \dots, q\}$

With respect to the parametrized portfolio model, we employ the original HSJ technique (E Downey Brill, Chang, and Hopkins 1982) with decision variable weighting to pursue ${}^{\text{PWA}}D$ diversity within a parameter ρ. We assume the ρ^{th} parameter also defines $\{(\rho, v, 1,0) \mid \forall v \in V_\rho\}$ technical constraints (i.e., $\{(\rho, v, 1,0) \mid \forall v \in V_\rho\} \subseteq L$). For example, the ρ^{th} parameter defines a project while the rest of the element parameters defines a project implementation mode (i.e., how the project is executed or deployed). We replace the objective function of the original optimization model with equation (25) in the HSJ adjusted model that minimizes the weighted sum of elements that possess parameter values equal to previously selected elements. Constraint (26) represents HSJ's near-optimality constraint. It is used to restrict alternative generation to near-optimal solutions. We slightly deviate from HSJ, as proposed by Brill, to accommodate the decision maker requesting a specific number of unique solutions, n, in the solution set by adding constraints (27). These constraints represent a type of Gomory cut (Gomory 1958); they ensure no identical solutions, as defined by the elements' ρ parameter, are added to the solution set more than once. Voll et al. (2015) provide an alternative integer cut formulation that could replace constraint set (22) for problems without a fixed number of selections constraint (22).

$$\min \sum_{v \in V_\rho} \sum_{i \in I_{\rho=v}} w_{\rho,v} x_i \tag{25}$$

subject to

Constraints (21), (22), (23), (24)

$$\sum_{i=1}^{q} b_i x_i \geq T \tag{26}$$

$$\sum_{j \in S_k^1} \sum_{i \in I_{\rho=j_\rho}} x_i \leq (m-1), \qquad \forall k \in \{1, \ldots, \alpha\} \tag{27}$$

where α denotes the number of solutions in S ($\alpha \leq n$), $w_{\rho,v} = \sum_{k=1}^{\alpha} \sum_{j \in I_{\rho=v}} S_{k,j}$ (i.e., count of selected elements in solutions of S that have the ρ^{th} parameter equal to v), and $S_k^1 = \{i | S_{k,i} = 1\}$.

To address the low correlation preference of solution set diversity, we propose an extension to the HSJ technique which we denote as 'HSJP' (Hop Skip Jump Pairs). At each HSJ iteration, the adjusted model incorporates additional binary decision variables, $y_{g,h}$ indicating the decision to select two elements that have the ρ^{th} parameter equal to g for one element and equal to h for another element, for each selection pair combination in the previous solutions. Constraints (29) and (30) control the selection of these auxiliary decision variables and forces the selection of the auxiliary decision variables. We incorporate these decision variables as part of the HSJ adjusted model's objective function, equation (28), to minimize the reselection of element combinations (i.e., the correlation of the element selections in the solutions). This causes the adjusted model to penalize the selection of element pairs and related element pairs if they were already selected together in past solutions.

$$\min \sum_{v \in V_\rho} \sum_{i \in I_{\rho=v}} w_{\rho,v} x_i + \sum_{(g,h) \in P_S} w_{\rho,(g,h)} y_{g,h} \tag{28}$$

subject to

Constraints (21), (22), (23), (24), (26), (27)

$$\sum_{i \in I_{\rho=g}} x_i + \sum_{i \in I_{\rho=h}} x_i - 2y_{g,h} \leq 0, \qquad \forall (g,h) \in P_S \tag{29}$$

$$y_{g,h} \in \{0,1\}, \qquad \forall (g,h) \in P_S \tag{30}$$

where $w_{\rho,(g,h)} = \sum_{k=1}^{\alpha} (\sum_{i \in I_{\rho=g}} S_{k,i})(\sum_{j \in I_{\rho=h}} S_{k,j})$ and $P_S = \{\text{distinct } (g,h) | g = i_\rho, i \in S_k^1, h = j_\rho, j \in S_k^1, i \neq j, \forall k \in \{1, \dots, \alpha\}\}$ (i.e., the set of unique ρ^{th} parameter value pairs within a solution of S).

5.3 Application problem

Fantasy sport participants face the portfolio problem in many types of fantasy sport competitions when the players take the form of the selectable elements; this paper refers to the portfolio problem composed of player selection decisions as the lineup selection problem. The lineup selection problem of interest is based on a decision maker facing the choice of which players to select to maximize the points realized from the selected players in the next day's game performances (with the points computed from their performances in these games). For each day of games, the participant is presented a decision problem in which they must select 8 players from any of the players taking part in the next day's games with technical constraints restricting the number of players for a certain position and a budget constraint.

The lineup selection problem of interest is manifested as part of a fantasy sports DSS. The DSS receives player performance predictions from an external proprietary model. The DSS possesses some means to tailor the mathematical model to users' unique preferences; the DSS can only accommodate changes to the model that are anticipated. For example, Smith, Sharma, & Hooper (2006) discover personalized heuristics fantasy sports participants utilize to support lineup decisions. The lineup selection optimization incorporates the player performance predictions and any custom decision maker lineup preferences into a binary linear programming model to find an optimum lineup that maximizes the total number of points from the lineup's selected players. Considering a real-world dataset from this DSS consisting of 5 games and 147 decision variables, this problem instance has 1,433 near-optimum solutions 2.5% from the optimal solution value and 111,550 near-optimum solutions 5% from the optimum solution value. The MGA technique is used by the DSS to generate a small set of these near-optimum solutions to present to the DSS user.

The lineup optimization model takes the form of the binary linear programming model presented in Section 2. The decision element tuples are formed from two parameters ($\delta = 2$). The first parameter, i_1, represents the player. The second parameter, i_2, represents the position (i.e., role). The possible positions include: point guard (PG), shooting guard (SG), center (C), small forward (SF), and power forward (PF). For example, the DSS user may choose a basketball player to fulfill the role of the PG or the role of the SG assuming the given basketball player fulfilled the PG and SG role in the past. The ρ parameter is set to 1 to base diversity on the player parameter of the selection elements. The original objective, equation (20), is to maximize the number of points from the player selections where b_i represents the predicted points the player of i is expected to generate. The decision maker is given a budget and player costs that constrain the lineup selection, constraint (21). The decision maker must select 8 players (i.e., $m = 8$). Constraints (23) are defined given

$$L = \{(2,\text{PG},1,1),(2,\text{SG},1,1),(2,\text{C},1,1),(2,\text{SF},1,1),(2,\text{PF},1,1),(2,\text{SF} \cup \text{PF},3,1),(2,\text{PG} \cup \text{SG},3,1)\} \cup M$$

where $M = \{(1, v, 1,0) | \forall v \in V_1\}$ (i.e., constraints to ensure a player is not selected for more than one position). The model also includes optional constraints (e.g., do not select players from a given team) or objective function embellishments (e.g., penalize the point estimates for a given set of players) made available to the fantasy sports player through the decision support system. HSJ and HSJP incorporates these optional embellishments by including the changes into the adjusted model as requested by the DSS user.

Regarding algorithm complexity, the original portfolio optimization model, the HSJ adjusted model, and the HSJP adjusted model are binary linear programming models. The iterative nature of the HSJ and HSJP techniques requires $n - 1$ executions of the optimization technique employed to solve the adjusted binary linear programming model. The complexity of the solving binary linear programming models can be analyzed through the number of decision variables and constraints (Meindl and Templ 2012). The original optimization model requires approximately 300 decision variables for problem sizes derived from 10 games (10 games, 2 teams per game, 12 players per team, 1.25 positions per player, i.e., on average 25% of players can be selected for 2 positions). The original HSJ technique adds one constraint for the near-optimality constraint and does not add any additional decision variables for the adjusted optimization model. HSJP may add up to 28 (8 choose 2) additional decision variables and constraints at each iteration for each player selection pair. Considering HSJP generating 50 additional solutions (i.e., 50 iterations of adding 28 decision variables and constraints), HSJP may increase the complexity of the binary programming model by up to 1400 additional decision variables and 1400 additional constraints. Meindl & Templ demonstrate the positive computational feasibility of integer linear programming formulations more complex than this in regard to the number of decision variables and constraints; Section 5 demonstrates the computational feasibility of the HSJP technique applied to the portfolio selection problem of interest. Equation (31) provides an upper bound to the number of additional decision variables and constraints HSJP may introduce for the nth iteration of the adjusted model where m represents the maximum number of elements in a portfolio (e.g., $m = 8$ for the applied lineup selection problem of interest).

$$\min\left(\binom{m}{2}(n), \binom{q}{2}\right) \tag{31}$$

5.4 Experiment

To evaluate the proposed HSJP technique's ability to increase diversity through the reduction of solution set correlation, an extensive numerical experiment is performed with 225 problem instances derived from the distribution of parameters held by a real-world basketball fantasy sports dataset. The real-world dataset consists of predicted points, salaries, and positions for players for five basketball games taking place the next day. 25 random problem instances are generated for problem sizes consisting of 2 to 10 games using player-parameter distributions discovered from the real-world dataset. For each game, 24 (2 teams of 12 players) players with salary, points, and position(s) parameters are randomly generated. An offset, exponential random variable distribution is used to generate the players' salaries. A linear relationship is observed in the real-world dataset between the players' salary and point variables. Linear regression analysis with salary as the input and the predicted points as the regressor results in a good fit. The standard deviation of the regression predicted values from actual values differences is used to generate a random normal distribution error from the regression predicted points given the salary random variable. Any generated random point value less than 0 points is assigned 0 points. Each player is randomly, uniformly assigned a position from the standard basketball positions. 25% of the players, roughly the percentage of players in the real-world dataset having two possible position options, are randomly, uniformly assigned a second position from the remaining positions.

The original HSJ technique and the proposed HSJP technique are executed to generate 50 solutions for each problem instance and for numerous near-optimality thresholds ($T = 2.5\%, T = 5\%, T = 10\%$, and $T = 15\%$ decrease from the initial solution's original objective value). The initial solution generated is the first optimal solution found to the original optimization formulation. The adjusted model is iteratively solved until the solution set contains 50 solutions or until the adjusted model is unable to find an additional unique near-optimal solution. CPLEX 12.7 optimization software is used, with default settings, on Intel Xeon processor computers to solve the binary linear programming model optimizations. The low correlation diversity objective is evaluated for solution sets sizes of 10, 25, and 50 solutions with respect to two measures, described directly subsequently, that negatively value solution sets that possess more element selection correlations: the Solaw-Polasky measure, ${}^{\mathrm{SP}}D$; and a measure this paper presents based on information entropy theory denoted entropy weighted variance, ${}^{\mathrm{I}}D$. To assess significance of diversity differences, paired t-test statistics (paired by problem instance) are computed, confirming paired t-test assumptions hold. Statistical significance is determined with a p-value being smaller than a 1% threshold.

5.4.1 Correlation sensitive diversity measurements

Solow & Polasky (1994) recognize that many diversity indexes and applications ignore the "distance" between objects in a set. They provide the example "a set consisting of four species of ants is some sense less diverse than a set consisting of one species of ant, one species of elephant, and one species of fern." They propose three requirements for a set diversity measure and called measures that met these criteria a "pure" diversity measure and propose equation (35) as a "pure" diversity measure. They show equation

(35) is a lower bound for a probability-based benefit (i.e., utility) function and propose an interpretation of the measure as the "effective number of species". A properly calibrated Solow-Polasky measure results in the value of 1 when each element of the set is significantly the same and up to a value of the cardinality of the set size n, the number of objects in the set, when every object in the set is completely unique in regard to the diversity benefit. A starting point for the Solow-Polasky measure are distance functions, denoted in regard to portfolios as $d(S_k, S_l)$, that quantify the distance between two portfolios of the solution set S. The $d(S_k, S_l)$ function by definition meets the conditions specified by equations (32), (33), and (34). The Solow & Polasky measure requires the specification of f, a positive definite function, that acts as a "correlation" transformation of the distance computations (f results in 0 if no correlation, 1 if completely correlated). In this experiment, three configurations of the Solow-Polasky measure ($\theta = 0.25$, $\theta = 0.5$, or $\theta = 0.75$) are considered using the pair-wise difference count as $d(S_k, S_l)$.

$$d(S_k, S_l) \geq 0 \tag{32}$$

$$d(S_k, S_k) = 0 \tag{33}$$

$$d(S_k, S_l) = d(S_l, S_k) \tag{34}$$

$$^{\mathrm{SP}}D(S) = \mathbf{1}'F^{-1}\mathbf{1} \tag{35}$$

$$F = \begin{bmatrix} f(d(S_1, S_1)) & \cdots & f(d(S_1, S_n)) \\ \vdots & \ddots & \vdots \\ f(d(S_n, S_1)) & \cdots & f(d(S_n, S_n)) \end{bmatrix} \tag{36}$$

$$\mathbf{1} = n \text{ vector of 1s} \tag{37}$$

$$f(d) = e^{-\theta d} \tag{38}$$

The diversity objective is also evaluated with a proposed measure that summarizes correlation insight from a solution set's covariance matrix with information entropy concepts, which this paper denotes as entropy weighted variance, $^{\mathrm{I}}D$. Similar to how principal component analysis uses eigenvectors of the covariance matrix to reduce the number of dimensions in the original dataset (Dillon and Goldstein 1982), the average amount of information needed to encode the solutions differences in these hidden components is estimated.

The player/project pairwise selection covariance matrix, equation (18), is a summary of how each selectable player of a solution set varies from the portion of times it was selected (i.e., mean) in relation to the other selectable players of the dataset. If two selectable players vary from their mean in opposing (similar) directions, then the covariance between the two selectable players is negative (positive). Likewise, if two selectable players vary from the mean randomly or out of sync with respect to one another the covariance tends to zero. Eigenvectors of the covariance matrix are the vectors, w, that satisfy equation (39) where $\boldsymbol{\lambda}$ represents the set of eigenvalues of $\text{cov}(S)$. Since the covariance matrix is positive semi-define, the eigenvalues are non-negative (Horn and Johnson 1985). The eigenvectors of the covariance matrix represent a set of orthogonal directions that explain the variance within the original dataset. The direction of the eigenvectors account for how the dataset's selected players are included in lineup solutions together. The eigenvector corresponding to the largest eigenvalue represents the transformed dimension that has the strongest correlation in the dataset.

$$\text{cov}(S)w = \boldsymbol{\lambda} w \tag{39}$$

The reduced, orthogonal dimensions (i.e., the components) are assumed to represent hidden, underlying factors of the variance observed in the dataset. Given that the sum of the covariance matrix eigenvalues equals the sum of variance and that the eigenvalues represent the amount of variance explained by the corresponding eigenvector-based component, the portions of the total variance explained by each underlying component can be determined. For example, if the first two largest eigenvalues sum to a value 80% of the total variance, the transformed components represent or retain 80% of the total variance in the dataset. Considering every difference unit being linked to an underlying-hidden component, the portion of these difference units linked to an underlying component are represented by the normalized eigenvalues. Let $\boldsymbol{\lambda}^{\mathrm{n}}$ represent the non-zero, normalized vector of eigenvalues of $\mathrm{cov}(S)$ (i.e., each eigenvalues is scaled by 1 over the eigenvalues sum (i.e., the total variance of S). Assuming the difference units are communicated or focused upon randomly and independently, the amount of information needed to communicate the specific source of all the variance units in a given dataset from the underlying component perspective is quantified using information entropy. The total variance is analogous to the expected number of messages needed to explain one of the solution's differences from the solution set mean. A term this paper denotes as covariance entropy, equation (40), is analogous to the expected size of a message distinguishing the underlying source component of a difference. Multiplying both results in the expected amount of information needed to distinguish the sources of the differences from the dataset's mean of a single lineup within a solution set (in bits given $\log_2$). This paper denotes equation (41) as the entropy weighted variance.

$$H(\boldsymbol{\lambda}) = -\sum_{i}^{|\boldsymbol{\lambda}|} \lambda_i^{\mathrm{n}} \log_2(\lambda_i^{\mathrm{n}}) \tag{40}$$

$$^{\mathrm{I}}D(S) = \left(\sum_{i}^{|\lambda|} \lambda_i\right) H(\boldsymbol{\lambda}) = \frac{^{\mathrm{PWA}}D(S)}{2} H(\boldsymbol{\lambda}) \tag{41}$$

The results are also compared using the pair-wise difference average diversity measure, ${}^{\text{PWA}}D$, however this common measurement of diversity does not take into account the correlation of the solution set's selections.

5.5 Results and Discussion

Regarding diversity of the generated solutions, Table 5 shows the paired t-test results, paired by problem instance, comparing the solution set diversity measure difference from the original HSJ method to the proposed HSJP method for numerous parameter configurations and diversity measures. '+' denotes the proposed HSJP is significantly better (i.e., HSJP produces more diverse solution sets given a 0.01 significance level) than the original HSJ technique, '-' denotes the original HSJ technique is significantly better, and '=' denotes failure to find any difference. For every correlation- sensitive diversity measure and for every configuration with the exception of two test comparisons at one configuration setting, the proposed HSJP technique produces significantly more diverse solution sets. The two test comparisons where the difference is not significantly different, occurs when the solution set size in small ($n = 10$) and the near-optimal criteria is at the largest considered threshold ($T = 15\%$). This suggests if the near-optimality threshold enables many solutions and only a small number of decision-space-diverse solutions are requested, the original HSJ technique may result in solution sets with low-correlation matching that of HSJP. With respect to diversity using ${}^{\text{PWA}}D$, the paired t-test suggests no significant diversity difference exists between HSJ and HSJP for 3 of the 12 parameter configurations. At the remaining 9 parameters configurations, the original HSJ method's solution sets are more diverse in regard to the correlation in-sensitive ${}^{\mathrm{PWA}}D$ measure.

TABLE 5. PAIRED DIVERSITY DIFFERENCE T-TEST SIGNIFICANCE RESULTS

T	2.5%			5%			10%			15%		
Metric \ n	10	25	50	10	25	50	10	25	50	10	25	50
${}^{\mathrm{PWA}}D(S)$	-	=	-	-	=	-	-	-	-	-	=	-
${}_{\theta=0.25}^{\mathrm{SP}}D(S)$	+	+	+	+	+	+	+	+	+	=	+	+
${}_{\theta=0.5}^{\mathrm{SP}}D(S)$	+	+	+	+	+	+	+	+	+	+	+	+
${}_{\theta=0.75}^{\mathrm{SP}}D(S)$	+	+	+	+	+	+	+	+	+	=	+	+
${}^{\mathrm{I}}D(S)$	+	+	+	+	+	+	+	+	+	=	+	+

Figure 17 shows the diversity t-test difference comparisons by problem size, diversity measure, and near-optimality threshold (each comparison incorporates 25 of the 225 problem instances). For every correlation sensitive diversity comparison, HSJP generates an equal or significantly more diverse solution set compared to HSJ. At $n = 25$ and $n = 50$, the number of comparisons where HSJP generates significant more diverse solution sets compared to HSJ is greater than the number of similar comparison results at $n = 10$. This suggests HSJP's potentially greater usefulness for uses when $n > 10$ for the application problem. At $n = 50$, $T = 2.5\%$, and for four of the five smallest problem sizes, the ${}_{\theta=0.25}^{\mathrm{SP}}D(S)$ measure suggests the diversity of the HSJP's solution sets are equal to HSJ's solution sets. This suggests the solution sets may be approaching covariance limits given the finite number of near-optimal solutions and some curvature in the amount of covariance reduction possible given the restrictive near-optimality threshold and the smaller problem sizes. At $n = 25, 50$ and $T = 10\%, 15\%$, HSJP generates more significant diverse results in regard to ${}^{\mathrm{PWA}}D(S)$. This demonstrates that the greedy heuristic methods of the original HSJ technique at times fails to minimize ${}^{\mathrm{PWA}}D(S)$. Note HSJP retains the greedy heuristic attributes to pursue correlation-sensitive diversity.

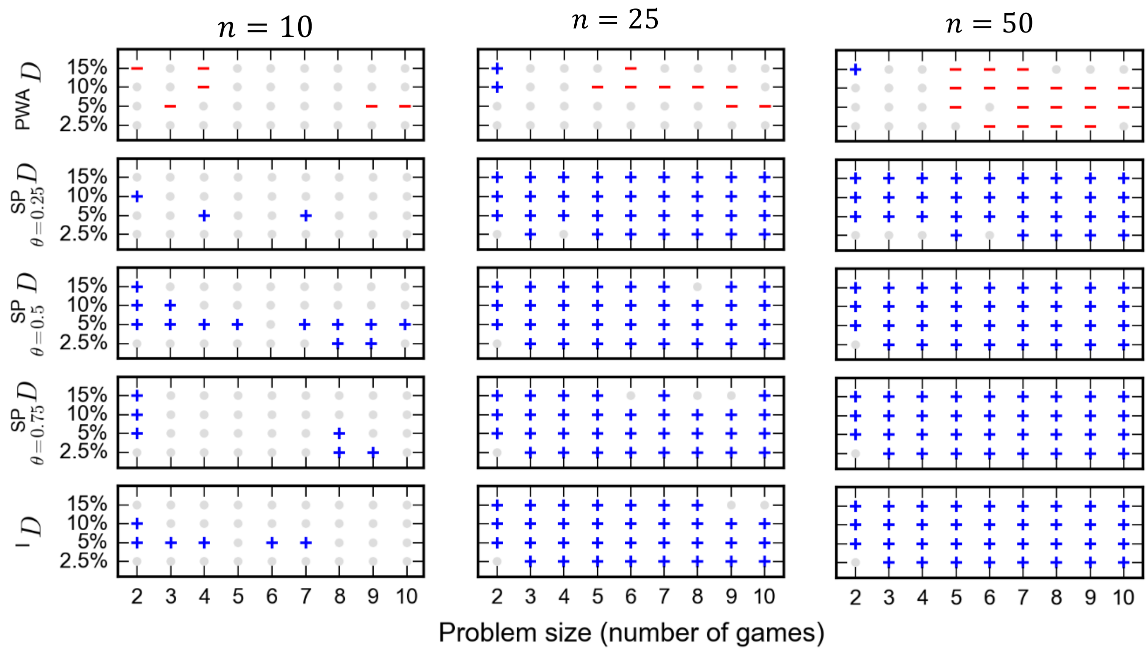


FIGURE 17. PAIRED DIVERSITY DIFFERENCE T-TEST RESULTS FOR DIFFERENT PROBLEM SIZES BY SOLUTION SET SIZE (n)

Figure 18 summarizes the amount of covariance in the generated solution sets for both the orignal HSJ technique and the proposed HSJP technique by taking the average $2^{H(\lambda)}$ (note that $H(\lambda)$ is the entropy of the solution sets' covariance matrix eigenvalues) over the 225 problem instances by each near-optimiatliy threshold value (T). $2^{H(\lambda)}$ is anagolous to the number of equally important components needed to explain the variance in the solution set. The more number of components required to explain the variance (higher entropy) the less solution set correlation. HSJP at $T = 10\%$ almost exceeds the covariance entropy of HSJ at $T = 15\%$ highlighting the ability of HSJP to reduce correlation in solution sets without requring the expanision of the near-optimalitiy criteria. As the solution set size (n) increases to 50 at the less restrictive near-optimaility thresholds, HSJP's ability to generate low-correlated solutions increases relative to HSJ compared to the most restrictive near-optimiatily threshold of $T = 2.5\%$. Figure 19 also shows this difference increase. Note that increasing the solution set size equal to the number of near-optimal solutions causes the same solution set from each MGA technique and subsquently the same diversity. As n increases and approaches the number of near-optimal solutions, this diversity difference is expected to approach zero.

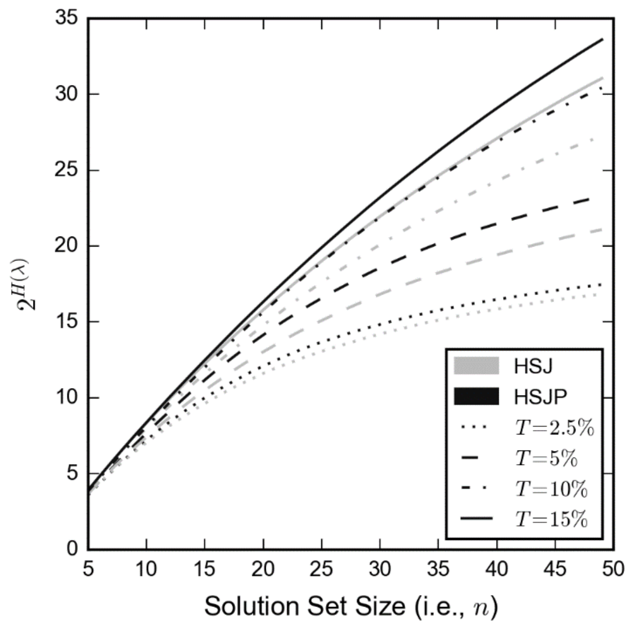


FIGURE 18. SOLUTION SET COVARIANCE ENTROPY AVERAGES

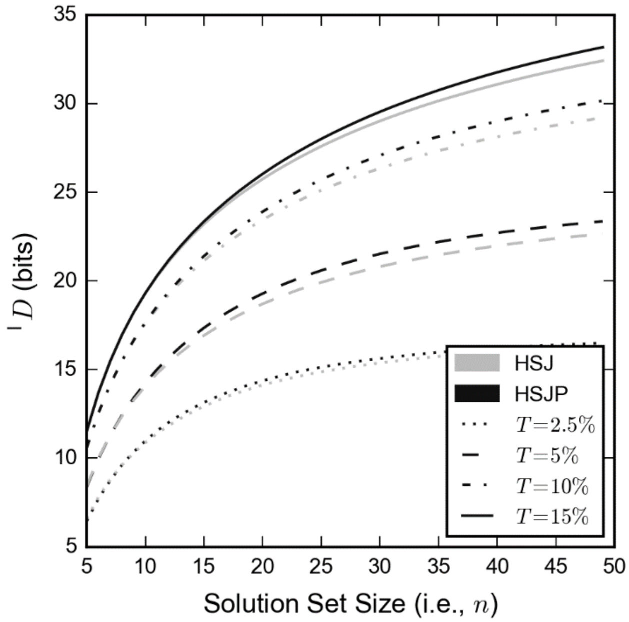


FIGURE 19. SOLUTION SET DIVERSITY AVERAGES

For every problem instance HSJ and HSJP generate 50 near-optimal solutions in acceptable times for the decision support system. The computational times range from 10 seconds to approximately 4 minutes. Figure 20 shows the minimum, maximum, median, 25% quartile, and 75% quartile computation times for the 225 problem instances. HSJP demonstrates more varied and longer computational times given the additional decision variables and constraints that increase the problem size modeled in the adjusted model's binary linear programming formulation.

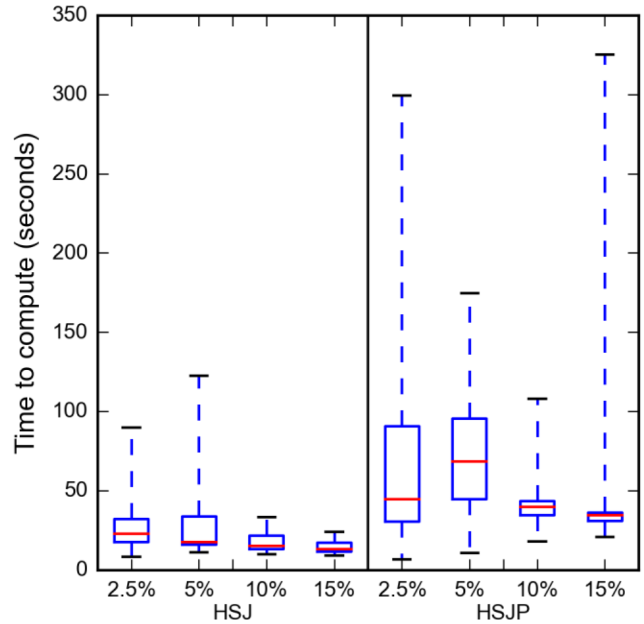


FIGURE 20. COMPUTATIONAL TIMES TO GENERATE 50 SOLUTIONS BY NEAR-OPTIMALITY THRESHOLDS

Finally, note HSJP retains the positive features of HSJ for use in a decision support system. HSJP can be applied to any decision support system modified linear programming optimization formulation without extensive calibration, such as the binary linear programming formulation of the lineup selection problem presented in Section 2. Like HSJ, HSJP only requires a threshold for near-optimality input to extend the original optimization model. Also, decision support systems (DSS) utilizing HSJP can generate and present additional solutions in a piecemeal fashion as desired by the DSS user.

5.6 Conclusions

This paper revisits modeling to generate alternatives (MGA) with respect to a portfolio selection problem with a binary linear programming optimization formulation. This paper presents a new HSJ technique variation that incorporates minimizing the selection of past selected pairs to produce multiple, low-correlated diverse near-optimal solutions. The proposed HSJP technique is compared to a variant of the original HSJ technique regarding the low-correlation diversity objective of the MGA techniques' solution sets. The results of a numerical experiment show the proposed HSJP technique produces significantly more diverse solution sets for 45 of the 48 comparisons given different near-optimality thresholds, different solution set sizes, and four different correlation-sensitive diversity measures.

Given the results, several future issues may benefit from more research. Future research is recommended to explore the genetic and evolutionary algorithm based MGA techniques compared to the proposed HSJP technique with respect to the correlation sensitive diversity measures. Screening MGA techniques that account for selection correlations are another suggested research avenue to improve MGA techniques with respect to correlation sensitive diversity. Future research is also suggested into the robustness and the calibration of MGA techniques for use in a daily-used, multi-user DSS considering daily changing inputs, specifically methods that automatically calibrate and adjust the near-optimality threshold requirements for techniques such as HSJ and HSJP.

VI. Progress towards addressing multiple, non-constant marginal value objectives

6.1 Introduction

Past multi-objective approaches applied to the project selection problem suggests summing the normalized project values computed from multi-attribute measurable value (Golabi, Kirkwood, and Sicherman 1981). An assumption of the multi-attribute measurable value approach is that the decision maker's portfolio preferences have constant marginal values with respect to the objective's measurement criteria. For some project selection problems the constant marginal value assumption does not hold (Kleinmuntz 2007). For example, consider a city management organization that is considering a number of projects to improve the city's environment. Assume the city holds two objectives of maximizing tree-park space and maximizing drinking water quality. If a portfolio contains a set of projects that as a group result in great water quality, the city organization may value a project that contributes additional water quality less than valuing the same project with respect to a portfolio that contains a set of projects that results in low water quality as a whole. Economists refer to this common preference phenomenon as the law of diminishing marginal utility.

This paper presents an approach to support optimization of the multi-objective project selection problem in regard to the objectives' measurement criteria having non-constant marginal value. The approach permits the decision maker to holistically evaluate the objectives of a portfolio, while also leveraging common multi-objective decision analysis value elicitation techniques as suggested by (Keeney and Raiffa 1976). The project selection problem is formulated as a non-linear binary integer programming model. Specifically, the formulation could be described as a binary integer version of a sigmoidal programming problem (Udell and Boyd 2013).

Next, Section 6.2 presents a literature review of techniques to address this issue. Section 6.3 proposes a new methodology to address this issue and an optimization solver to optimization the mathematical model. Section 6.4 provides the results of a demonstration of the methodology in regard to an environmental project selection problem compared to two other optimization approaches.

6.2 Literature Review

Past multi-objective project selection research works around optimization modeling limitations that force constant marginal value assumptions with iterative decision maker methods. Golabi et al. (1981) employ an iterative method that solves the linear programming optimization model, presents the solution to the decision maker to evaluate, and allows the decision maker to constrain objectives to be above a chosen level with a linear constraint to explore solutions that may capture relatively higher marginal value, which the original binary linear programming model's optimization did not discover. Dickinson et al. (2001) exclude some objectives from the objective function and modeled the objectives as constraints. This paper denotes techniques that enforce the realization of some non-constant marginal value of an objective with a constraint as a constraint estimator. This paper denotes techniques that ignore non-constant marginal values to model portfolio objectives in a linear objective function as a linear estimator. Figure 21 shows the potential effects of constraint estimations and linear estimations on representing non-constant marginal value objectives where V_o denotes the portfolio value function for objective o and z_o denotes objective o's raw criteria measurement contributed by a set of selected projects. Constraint estimators only permit solutions that are equal or greater than the constraint. In this manner, constraint estimators ignore or undervalue solutions that fall below the constraint and overvalue solutions that barely meet the constraint since these solutions are valued the same as a solution with the best possible value on the criterion range. Linear estimators over value solutions when the solution results in a value that is larger in regard to the true decision maker value curve. Kleinmuntz (2007) describes a common workaround used in the application of portfolio decision analysis problems denoted "threshold constraints" that uses both, a constraint estimator and a linear estimator.

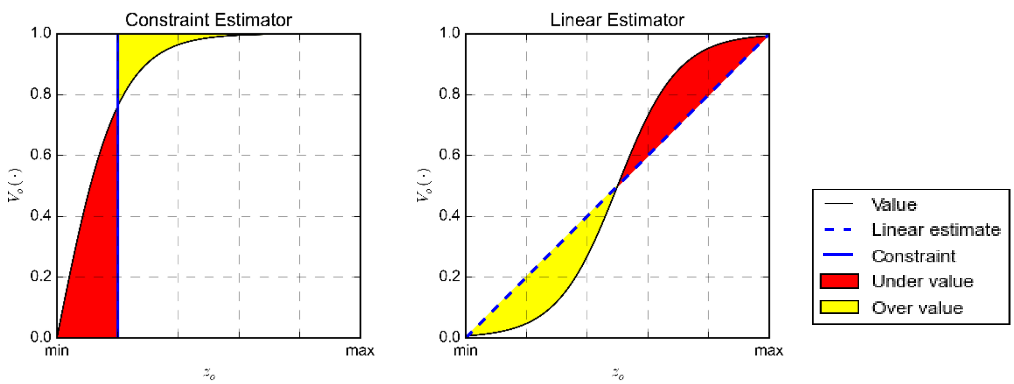


FIGURE 21. ESTIMATION EFFECTS ON NON-CONSTANT VALUES

Numerous methods to address projects' benefit interactions have been proposed. Table 6 provides a summary of past approaches discovered to address benefit interactions between projects in multi-objective project selection problems. A small set of this research (Argyris, Figueira, and Morton 2011; Liesiö 2014) directly addresses the non-constant marginal value benefit interaction type. A limitation of the approach suggested by (Liesiö 2014), as it was presented, is that it relies on unproven project-centric decision maker elicitation methods to construct symmetric multilinear value functions that may be unnatural for decision maker preference modeling with non-constant marginal values and may be inappropriate when projects may largely vary in scope, size, and objective significance. For example, if one project being considered contributes a very significant amount to one objective many multiples more than any other project, the project centric value function would have to reflect this for this one project. Subsequently, eliciting for the objective's portfolio value function for this criteria, as proposed by (Liesiö 2014), requires the use of the best project with respect to this criteria as this elicitation step-size; the large step-size caused from this one project can gloss over the smaller non-constant marginal preference changes through the objective criteria range. A limitation of the method proposed by (Argyris, Figueira, and Morton 2011) is that the value functions are restricted to concave or linear forms (i.e., convex marginal values are not supported). In the next section, the paper proposes a non-linear optimization methodology that addresses multi-objectives with concave and convex non-constant marginal value attributes while reducing the need to use project-centric value elicitation methods to formulate the optimization model's objective function.

TABLE 6. PROJECT BENEFIT INTERACTION RESEARCH

Research	# of Objectives	Objective Function Form	Solution Technique
(Carazo et al. 2010)	2 ,4 ,6+	Polynomial	Pareto-optimal, search
(Bhattacharyya, Kumar, and Kar 2011)	3+	Polynomial	Pareto-optimal, genetic algorithm
(Crama and Schyns 2001)	2	Polynomial	Simulated annealing
(Rabbani, Aramoon Bajestani, and Baharian Khoshkhou 2010)	3	-	Particle Swarm, SPEAII
(Dickinson, Thornton, and Graves 2001)	1	Polynomial	COTS non-linear solvers
(Blecic, Cecchini, and Trunfio 2012)	1	Polynomial	Search Heuristics, Evolution Strategy
(Eilat, Golany, and Shtub 2006)	1	Ratios	Data Envelopment Analysis
(Fox, Baker, and Bryant 1984)	1	Quadratic Polynomial	Linearization
(Stummer and Heidenberger 2003)	k	Polynomial	Linearization
(Liesiö 2014)	k	Symmetric Multilinear	Enumeration, Heuristic
(Argyris, Figueira, and Morton 2011)	k	Concave piece-wise linear additive	

6.3 Methodology

The methodology begins with an analysis of the portfolio's multiple objectives from a holistic portfolio perspective. At this point, the fact that portfolio-alternatives are defined by a combination of projects is ignored. This approach employs problem objective modeling and validation to ensure the objective attributes are mutually preferentially independent. This enables an additive value function representation of the decision maker preferences (Keeney and Raiffa 1976). This is the first requirement of the methodology. Next, value functions are constructed for each objective and the objectives are assigned weights using traditional multi-objective techniques with two exceptions. First, for each portfolio objective, value function inputs are to be based on extensive (i.e., addable) attributes (Krantz et al. 1971) computable for each project. This is the second requirement of the proposed methodology. In many instances, a natural extensive attribute exists to measure the contribution of a project to the portfolio objective, such as acres of land preserved in an environmental-preservation project selection problem. In other instances, a measurable value function with baseline definition may be required to translate project attributes into an extensive attribute, such as attributes specified on a Likert scale (from a SME-provided-assessment) in regard to the projects' contribution to the portfolio objective to improve city aesthetics. Second, each portfolio-objective's value function denoted V_o is a sigmoidal function. A sigmoidal function is defined as a function that is Lipschitz continuous and is either concave, convex, or concave to convex (or convex to concave) at a single point through the input range (Udell and Boyd 2013). Let O represent the set of objectives. Let P denote the set of potential projects with p denoting a project in this set. Let b_{po} represent p's measurement in regard to portfolio objective's o extensive attribute. Let z_o represents the sum of the selected projects extensive measurement regarding objective o and the input into V_o. Employing traditional weight elicitation methods to normalize the portfolio criteria values, let w_o denote the weight of objective o.

The resulting non-linear binary integer sigmoidal programming problem is represented by equations (42), (43), and (44) where x_p denotes the binary (i.e., yes/no) decision to pursue project p, R represents the set of limited resources, r represents one of these resources, a_r represents the amount of resource r available, and c_{pr} represents the resource cost to select project p with respect to resource r. Equation (44) represents the binary nature of the decision variables. A non-linear optimization solver is then utilized to prescribe a solution.

$$\max \sum_{o \in O} w_o V_o(\sum_{p \in P} b_{po} x_p) \tag{42}$$

subject to:

$$\sum_{p \in P} c_{pr} x_p \leq a_r, \forall r \in R \tag{43}$$

$$x_p \in \{0,1\}, \forall p \in P \tag{44}$$

To solve this problem, this paper extends the sigmoidal programming solver proposed by (Udell and Boyd 2013) to account for binary decision variables. The extended solver employs branch and bound techniques to address the binary decision variable characteristics. To seed the lower bound of the brand and bound technique, the proposed solver finds the optimization solution to the linear approximation of the problem. To generate the linear problem approximation, the solver replaces the non-linear value functions with a linear value function derived from the range of the non-linear value function inputs and anchoring the endpoints at the minimum and maximum values. Next, the branch and bound technique initially treats every binary decision variable as continuous and constrained between 0 and 1. The proposed solver branches on each decision variable to restrict the decision variable back to a binary 0 and a binary 1 value. The solver first proceeds in depth first manner exploring the binary 1 branch. It ensures the problem is feasibility and solves the continuous sigmoidal programming problem using Udell and Boyd's solver. If the continuous solution is less than the bound then the branch is pruned otherwise the branch is branched again. If the problem is at a leaf (all decision variables have been restricted to 1 or 0) and the value is more than the lower bound, then the lower bound is updated to this solution's optimal value and this solution is retained. The branch and bound technique enumerates every branch until every branched is pruned or evaluated.

6.4 Demonstration

To demonstrate the computational feasibility of the methodology, this paper considers an example and dataset studied and provided by (Liesiö 2014). The dataset consists of fifty projects representing unique areas to be conserved, with five decision maker objectives. The projects would publicly fund the purchase of privately-owned forests for a period of 10-20 years; public officials seek to maximize the conservation value of the selected portfolio. Portfolio value functions are reproduced after the multilinear value functions used by (Liesiö 2014). The reproduced portfolio value functions are not exact replicates given the different approaches to modeling the preferences; this precludes direct comparison of this paper results to the symmetrical multilinear optimized results. Three of the five objectives are naturally measured with an extensive attribute and do not require any additional project-centric value function. The objective to maximize area to be conserved is measured by the extensive (addable) acre property. The objective to maximize the protection of endangered species is measured by the extensive number of endangered animals protected property. The objective to maximize the amount of old broad-leaved trees is measured by the extensive volume (m^3) of trees. For the two remaining objectives, project-based measurable value function that converts the qualitative or quantitative project measurement into an extensive value for input into the portfolio value function are developed. The portfolio value functions are shown in Figure 22 and derived from equation (45) and the parameters in Table 7. The objective weights, resource needs, and resource limits are used as specified in the symmetrical multi-linear example.

$$V(z) = A + \frac{K - A}{1 + e^{(-B(z-M))}} \tag{45}$$

TABLE 7. OBJECTIVE VALUE FUNCTION PARAMETERS AND WEIGHTS

Objective	Objective Index	A	K	B	M	w_o
Maximum area	1	-0.3273	1.3273	0.02	70	0.15
Maximum trees	2	-1	1	0.002	0	0.1
Maximum Water Economy	3	-1.03	1.03	0.05	0	0.15
Maximum Homes of Endangered Species	4	0	1	0.007	750	0.25
Close to natural reserves	5	0	2	0.15	50	0.35

To study the effectiveness of the proposed solver, an off-the-shelf commercial non-linear solver provided with Microsoft Excel is used to solve the non-linear mathematical model. The proposed solver's solution is also compared to a solution from solving a linear relaxed variation of the model (the same one to seed the proposed solver). For the linear relaxed variation, a linear function over the range of the raw criteria range is mapped to the portfolio value from 0 to 1.

Figure 22 shows the solution's objective values to each optimization technique. The final value of the sigmoidal branch-and-bound non-linear solver solution is 0.5740, Microsoft Excel solver methods results a solution having a value 0.5737, while the linear relaxed method results in a solution having a final value of 0.5713 as computed from the non-linear model's objective function. While the proposed solver provides a better solution, uncertainties persist regarding the practical significance. Both non-linear solver solutions select 32 projects holding 2 project selection differences. Comparing the non-linear solutions to the linear relaxed solution, there are 7 and 8 different project-selection differences.

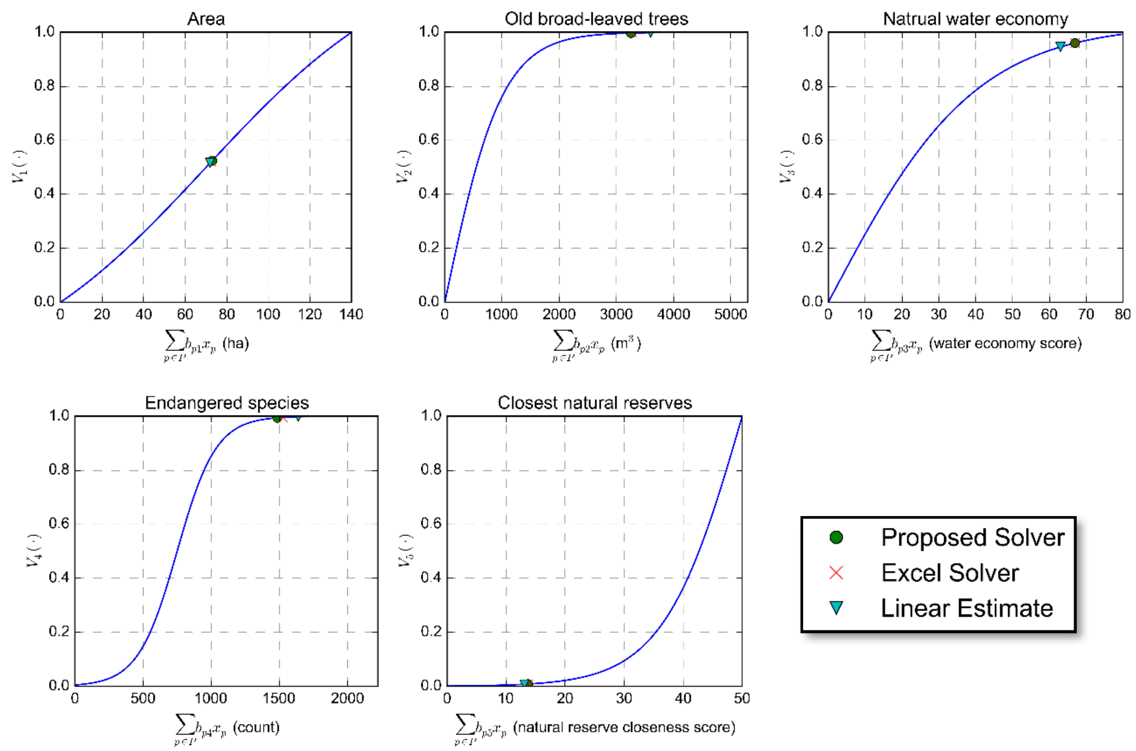


FIGURE 22. EXAMPLE PORTFOLIO VALUE FUNCTIONS AND SOLUTIONS

The approach in this paper is different than (Liesiö 2014) with respect to the method of formulating and eliciting the portfolio value functions. Liesiö's proposed method utilizes a symmetric multilinear preference formulation. The symmetric multilinear formulation, as presented, incorporates a project specific value function for each objective and anchors the elicitation of the criterion-specific portfolio value functions to the project with the most value for the criterion; they do not suggest a means to address non-constant marginal value less than this value. In doing so, the symmetric multilinear approach, as proposed, possibly requires more and unnatural elicitations. The proposed approach only needs a project specific value function if the portfolio objective's criteria does not possess natural ratio scale properties.

In addition, the proposed method of portfolio selection can be extended to support project interaction effects with regard to portfolio specific criteria. If a set projects hold an interaction effect with regard to a portfolio specific criterion, one could create an activation variable to account for this interaction effect and place a linear constraint on the project's decision variable to be less than or equal to both of the projects contributing to the interaction effect. One such extended model formulation is shown below where x_i denotes the activation variable for interaction i, b_{io} denotes the interaction's effect on objective o's criteria, P_i denotes the set of projects required to activate i, and I represents the set of interactions. Many other model embellishments are possible to account for different types of project benefit interactions; for example see (Stummer and Heidenberger 2003).

$$\max \sum_{o \in O} w_o V_o \left(\sum_{p \in P} b_{po} x_p + \sum_{i \in I} b_{io} x_i \right) \tag{46}$$

subject to:

$$\sum_{p \in P} c_{pr} x_p \leq a_r \text{ , } \forall r \in R \tag{47}$$

$$P_i \subset P, \qquad \forall i \in I \tag{48}$$

$$\sum_{p \in P_i} x_p \geq |P_i| x_i, \forall i \in I \tag{49}$$

$$x_i \in \{0,1\}, \qquad \forall i \in I \tag{50}$$

$$x_p \in \{0,1\}, \forall p \in P \tag{51}$$

6.5 Conclusion

In this paper an approach to the project selection problem is proposed that enables a decision maker to value project portfolios from the portfolio-as-a-whole perspective. The approach provides a means to model non-constant marginal values in portfolio objectives and other value interactions. Also, the approach enables the use of traditional, multi-objective value function elicitation methods as proposed by (Keeney and Raiffa 1976). The method is demonstrated with a dataset from literature and solved using a proposed branch and bound implicit-enumeration non-linear solver that finds a better solution compared to Microsoft Excel's non-linear solver and compared to a portfolio linear additive model that assumes constant marginal value.

An area of possible future research is to study how the decision context affects a decision makers' ability to accurately formulate their preferences using symmetric multilinear project-centric methods as proposed by (Liesiö 2014) compared to the method suggested in this paper. Another area of future research is to explore extending symmetric multilinear value functions elicitation methods to account for non-constant marginal value smaller than the best project. Another area of future research is on new or existing non-linear solver techniques and theory for the integer sigmoidal programming formulation. The extended model shows promising aspects to explicitly represent interaction effects and non-constant marginal values. Future research studying the feasibility in employing the extended model to support the project selection problem with project-set-benefit interactions is recommended.

VII. Summary and Conclusions

This dissertation provides novel techniques and methods to support resource allocation for information production activities. Viewing the information production problem in relation to research concerning project selection problems, this dissertation highlights research gaps in project selection methods and presents results to show the significance of the proposed methods to address these gaps.

Chapter III presents a methodology to support optimization of a project selection and scheduling with variable intensity work problem. The methodology is compared to a previously proposed methodology in regard to a dataset consisting of 1,800 problem instances. In every problem instance the presented methodology produces an equal or better solution.

Chapter IV presents a project selection problem consisting of projects that produce information. A novel methodology to support solution optimization is presented to address project selection and scheduling with project product deterioration and reproduction decisions. The chapter provides a formulation to compute a measurable value baseline with respect to time interactions on value function evaluations. The case study results show successfully employing the methodology on an information production planning problem for an Air Force organization.

Chapter V presents a new variation of a project selection alternative generation technique embeddable in project selection decision support systems. The chapter develops and shows the proposed alternative generation technique generates significantly more decision-space-diverse alternative sets compared to the original alternative generation technique. The chapter also utilizes a new set diversity measure based on information entropy to quantify the decision-space diversity of an alternative set.

Chapter VI presents progress towards an optimization method for the project selection problem with multiple, non-constant-marginal-value objectives. The chapter proposes an optimization solver and shows the solver finds a better solution compared to a commercial-off-the-shelf non-linear solver for a problem dataset provided in literature. The methodology incorporates proven multi-objective modeling techniques addressing a limitation of a recently proposed method incorporating unproven techniques.

The methods this dissertation proposes assume deterministic input parameters. A possible future research topic is possible extensions of these methods to address stochastic input parameters or incomplete information assumptions. Another possible future research area is into methods to support both the non-constant marginal objective criteria issues researched in Chapter VI and the scheduling issues addressed in Chapter III and IV.

VIII. Appendix A. Measuring portfolio set diversity

To evaluate the decision-space-diversity of a set of portfolios, this paper introduces a diversity index based on information entropy and the portfolio set's selection covariance matrix and shows how the index addresses shortcomings held by previously proposed diversity indexes. This paper demonstrates the ability of the proposed index to quantify diversity considering the selection correlations within a portfolio set.

8.1 Introduction

The objectives of alternative generation techniques incorporated into a decision support system are to produce diverse solutions (e.g., more diversity between the suggested portfolios leads to more opportunities for a decision maker to recognize and ultimately address any un-modeled preferences or other technical interactions between the selected elements in a portfolio) and that these solutions are nearly-optimal from the perspective of the original optimization model. To evaluate the decision-space-diversity aspect of these solution sets, this paper reviews methods other fields have used to quantify the diversity of a set of things, identifies shortcomings in these methods, and proposes a new diversity index to account for the shortcomings identified. The paper demonstrates how the proposed diversity index addresses these shortcomings and then discusses some of the properties held by the proposed diversity index.

8.2 Literature

Numerous measurement indexes have been used to evaluate an alternative generation technique's ability to overcome modeling inadequacies by quantifying the diversity of the generated solution sets. E Downey Brill et al. (1982) suggest three measures to quantify a solution set's ability to overcome modeling inadequacies: 1) sum of pairwise absolute differences in decision variables; 2) number of non-basic variables introduced; and 3) solution set range in mean-objectives. Researchers since have used variations of sum of pairwise absolute decision variable differences in the evaluation of MGA techniques (Baugh Jr., Caldwell, and Brill Jr. 1997; S.-Y. Chang, Brill, and Hopkins 1982; Loughlin et al. 2001; Shir et al. 2010; Zechman and Ranjithan 2007). Shir et al. (2010) use a sum of pairwise differences normalized to the diameter of the decision space to quantify the differences in a solution set. Ulrich et al. (2010) note a limitation of the pairwise difference measurements approach to quantify diversity. In continuous decision spaces, pairwise-difference optimizations often result in a set of points that overlap at the decision corners points and do not span across the decision space. After a discussion of approaches other fields have used to quantify diversity and the limitations of these and pairwise difference measures from the perspective of the lineup decision's binary decision space, this paper revisits Ulrich et al.'s approach to quantify the diversity of a set of solutions.

In other fields, researchers have proposed several different methods to quantify differences. Shannon (1948) introduces information entropy as a means to quantity uncertainty in the flow of information within a communication systems. Information entropy is a measurement of the amount of information needed to represent variability in information communications. Shannon defines discrete entropy as equation (52) where n represents the number of discrete events, x_i represents the discrete event (i.e., the symbols that may appear over an information channel), $P(x_i)$ represents the probability of event x_i, and b represents the information encoding units (such as $b = 2$ for bits).

Entropy, as defined here, is the minimum expected amount of information needed to distinguish between the set of discrete events. If these discrete events happen in time independently, the total amount of expected information needed to encode the result of m number of events is $mH(x)$.

$$H(x) = -\sum_{i=1}^{n} P(x_i) \log_b P(x_i) \tag{52}$$

(Jost, 2006) discusses and compares efforts to measure diversity from an ecology perspective. Jost advocates for "true diversities" of the equation (53) form, proposed by Hill (1973), where p_i denotes the portion of species i in the total animal population. Jost states that the diversity index should double when the number of species (or discrete event types) double assuming equal species proportions. Lucas et al. (2017) calls this the replication principle. Hill (1973) shows the relationship between Shannon Entropy (52) and the diversities of the form ${}^qD$ is $\lim_{q \to 1} {}^qD = b^{H(x)}$.

$${}^qD = \left(\sum_{i=1}^{n} p_i^q\right)^{\frac{1}{1-q}} \tag{53}$$

Lucas et al. (2017) surveys diversity indexes used in ecology for comparative analysis of molecular datasets. They list essential criteria for their index and find the diversity indexes of the form ${}^qD$ met all their criteria if the choice of q is done deliberately considering the trade-off weighting given to rare and abundant entities. Mcdonald & Dimmick (2003) present thoughts on the concept and measurement of diversity from a network radio programming perspective. They empirically show a few of the diversity indexes, including Shannon's entropy, $H(x)$, are sensitive to the "richness" (number of unique types, denoted here as R) and the abundance proportion properties.

The true diversity measures as proposed by (M. O. Hill 1973) are not directly applicable given the lack of a clear distinction of a species in the lineup selection problem. For example, if each portfolio that is different is considered a "species" then all solution sets of the same cardinality would be equal in diversity (all solutions are different since at least one selection in a portfolio is different from another portfolio) no matter how many different selection elements were incorporated and the dispersion of the selection in the suggested portfolios. A more relevant application of the true diversity measurements is possible if each selected project is considered as a species. Let p_j represent the portion that the project j was selected defined by the count of the number of times project j was selected divided by the total number of project selections in the entire solution set.

Pairwise difference, ${}^qD$, or $H(x)$ diversity measurements (with the specification of p_i as defined above) fail to recognize correlations of selected elements between portfolios within a solution set. Consider the solution set as a 0-1 multi-variate matrix, denoted here as S where the rows (indexed with i) represent the suggested portfolios, the columns (indexed with j) represent the selectable projects, and the values denote the selection (1) or non-selection (0). Table 8 shows solution set examples and some past methods to quantify diversity where ${}^{\mathrm{v}}D(S)$ represents the total variance of a solution set, equation (54), ${}^{\mathrm{PWT}}D(S)$ represents the total number of pairwise differences in a solution set, equation (55). Equation (56) shows the relationship between the average pairwise differences denoted ${}^{\mathrm{PWA}}D(S)$, to total pairwise differences, and to the total variance. See section 8.5 for proof of this scaler relationship between the total variance to the average number of pairwise differences.

$$^{\mathrm{v}}D(S) = \mathrm{var}(S_{n\times c}) = \frac{1}{n-1}\sum_{j=1}^{c}\sum_{i=1}^{n}\left(M_{i,j} - \mu_j\right)^2 \tag{54}$$

$$^{\mathrm{PWT}}D(S_{n\times c}) = \sum_{j=1}^{c}((\sum_{i=1}^{n} S_{i,j})(\sum_{i=1}^{n} 1 - S_{i,j})) \tag{55}$$

$$^{\mathrm{PWA}}D(S) = \frac{^{\mathrm{PWT}}D(S)}{\binom{n}{2}} = 2\ ^{\mathrm{v}}D(S) \tag{56}$$

TABLE 8. SOLUTION SET EXAMPLES WITH TRADITIONAL DISTANCE DIVERSITY MEASUREMENTS

S	$^{\mathrm{PWT}}D$, $^{\mathrm{PWA}}D$	$^{\mathrm{v}}D$	# of Elements Incorporated that changed	Portions or Relative Abundances (p_j)	$^{q=0}D$	$^{q=1}D$	$^{q=2}D$
[1,1,1,0,0,0,0,0,0,0,0,0], [1,1,0,1,0,0,0,0,0,0,0,0], [1,1,0,0,1,0,0,0,0,0,0,0], [1,1,0,0,0,1,0,0,0,0,0,0]	12, 2	1	4	4/12, 4/12, 1/12, 1/12, 1/12, 1/12	6	4.76	4.0
[1,1,0,0,1,0,0,0,0,0,0,0], [1,1,0,0,0,1,0,0,0,0,0,0], [0,0,1,1,1,0,0,0,0,0,0,0], [0,0,1,1,0,1,0,0,0,0,0,0]	24, 4	2	6	2/12 (for all 6)	6	6.0	6.0
[1,1,0,0,0,1,0,0,0,0,0,0], [1,0,1,0,1,0,0,0,0,0,0,0], [0,1,0,1,1,0,0,0,0,0,0,0], [0,0,1,1,0,1,0,0,0,0,0,0]	24, 4	2	6	2/12 (for all 6)	6	6.0	6.0
[1,1,1,0,0,0,0,0,0,0,0,0], [0,0,0,1,1,1,0,0,0,0,0,0], [0,0,0,0,0,0,1,1,1,0,0,0], [0,0,0,0,0,0,0,0,0,1,1,1]	36, 6	3	12	1/12 (for all 12)	12	12.0	12.0

Notice that these diversity indexes ignore the relationship of how the selection of one element in the portfolios varies compares to another element. For example, consider the second solution set in the table. The first four element selection variables (columns) vary together; either the first two elements are selected, and the next two elements are not selected or vice versus. These four decision variables could be reduced to one binary variable that is a composite of the original four variables. This reduction would drastically reduce the "richness" of the solution set. If the two first columns represents elements from group-A and the next two columns represent elements from group-B then the second solution set suggests 2 portfolios with elements mostly from group-A and 2 portfolios with elements mostly from group-B. Notice the third solution set would suggest 1 portfolio with elements mostly from group-A, 1 portfolio with elements mostly from group-B, and 2 mixed element portfolios. None of the diversity measures in Table 8 recognize this reduced correlation between the second solution set and the third solution set which does not possess these strong element selection correlations.

Ulrich et al. (2010) suggest a measure that can account for this. They incorporate a variation of a measure suggested by Solow & Polasky (1994) to quantify the diversity of a MGA method's solution set. Ulrich et al. provide a visual justification of the Solow-Polasky measure for decision spaces with continuous decision variables. They compare the maximization of pairwise distances distributions vs Solow-Polasky based distributions. The points that maximize pairwise distance measures cluster on top of each other at the corner points of the space while the points that maximize the Solow-Polasky measure are relatively distributed evenly across the space.

Solow & Polasky (1994) recognize that many diversity indexes and applications ignore the "distance" between objects in a set. They provide the example "a set consisting of four species of ants is some sense less diverse than a set consisting of one species of ant, one species of elephant, and one species of fern." They propose three requirements for a set diversity measure and called measures that met this criterion a "pure" diversity measure. Weitzman (1992) is the first to suggest a "pure" diversity measure and to show the measure met the diversity measurement criteria as formally presented by (Solow and Polasky 1994). A starting point for the Weitzman measure and the Solow-Polasky measure are distance functions, denoted $d(i,j)$, that quantify the distance between two objects, denoted here as i and j, that may belong to a set, denoted here as S. The $d(i,j)$ function by definition meets the conditions specified by equations (32), (33), and (34). The Weitzman measure, equation (60), uses recursion and employs a dynamic programming method for computation. Solow & Polasky (1994) propose equation (35) that met this criteria for some formulations. A properly calibrated Solow-Polasky measure results in the value of 1 when each element of the set is significantly the same and up to a value of the cardinality of the set size n, the number of objects in the set, when every object in the set is completely unique. The Solow & Polasky measure requires the specification of f, a positive definite function, that acts as a "correlation" transformation of the distance computations (f results in 0 if no correlation, 1 if completely correlated).

$$d(i,j) \geq 0 \tag{57}$$

$$d(i,i) = 0 \tag{58}$$

$$d(i,j) = d(j,i) \tag{59}$$

$${}^{\mathrm{W}}D(S) = \max_{s_i \in S}\left({}^{\mathrm{W}}D(S - s_i) + d_S(s_i, S - s_i)\right) \tag{60}$$

$$d_S(s_0, S) = \min_{s_i \in S} d(s_0, s_i) \tag{61}$$

$${}^{\mathrm{SP}}D(S) = e' F^{-1} e \tag{62}$$

$$F = \begin{bmatrix} f(d(s_1, s_1)) & \cdots & f(d(s_1, s_n)) \\ \vdots & \ddots & \vdots \\ f(d(s_n, s_1)) & \cdots & f(d(s_n, s_n)) \end{bmatrix} \tag{63}$$

$$e = n \text{ vector of 1s} \tag{64}$$

$$f(d) = e^{-\theta d} \tag{65}$$

The "pure" diversity measurements demonstrate positive properties regarding the correlation of differences in the example solution sets, see Table 9. The pairwise count of differences (hamming distance) represents the $d(i,j)$ function for these measurements.

TABLE 9. SOLUTION SET EXAMPLES WITH "PURE" DIVERISTY MEASUREMENTS

Solution Set (S)	$d(i,j)$ Matrix	${}^{W}D$	${}^{SP}_{\theta=0.5}D$
[1,1,1,0,0,0,0,0,0,0,0,0], [1,1,0,1,0,0,0,0,0,0,0,0], [1,1,0,0,1,0,0,0,0,0,0,0], [1,1,0,0,0,1,0,0,0,0,0,0]	[[0. 2. 2. 2.] [2. 0. 2. 2.] [2. 2. 0. 2.] [2. 2. 2. 0.]]	6.0	1.90
[1,1,0,0,1,0,0,0,0,0,0,0], [1,1,0,0,0,1,0,0,0,0,0,0], [0,0,1,1,1,0,0,0,0,0,0,0], [0,0,1,1,0,1,0,0,0,0,0,0]	[[0. 2. 4. 6.] [2. 0. 6. 4.] [4. 6. 0. 2.] [6. 4. 2. 0.]]	10.0	2.58
[1,1,0,0,0,1,0,0,0,0,0,0], [1,0,1,0,1,0,0,0,0,0,0,0], [0,1,0,1,1,0,0,0,0,0,0,0], [0,0,1,1,0,1,0,0,0,0,0,0]	[[0. 4. 4. 4.] [4. 0. 4. 4.] [4. 4. 0. 4.] [4. 4. 4. 0.]]	12.0	2.84
[1,1,1,0,0,0,0,0,0,0,0,0], [0,0,0,1,1,1,0,0,0,0,0,0], [0,0,0,0,0,0,1,1,1,0,0,0], [0,0,0,0,0,0,0,0,0,1,1,1]	[[0. 6. 6. 6.] [6. 0. 6. 6.] [6. 6. 0. 6.] [6. 6. 6. 0.]]	18.0	3.48

A limitation of the Solow-Polasky measure is the requirement for the specification and calibration of an additional function to produce a relatable diversity index. If the f function of the Solow-Polasky measure is underestimated any difference looks significant and every solution in a set would look totally different reducing our ability to quantify any small diversity differences between the solutions. A limitation of the Weitzman (dynamic programming computed) measure is that it fails to compute in a reasonable time for solution set sizes greater than 20.

Section 8.3 introduces entropy weighted variance that: 1) is capable of quantifying solution set diversity in light of the element selection correlations between the portfolios using information entropy theory; 2) is computationally tractable; and 3) does not require the specification of an additional input from the user.

8.3 Diversity Measurement Methodology

To evaluate the diversity of the solution sets and to address the limitations of the diversity measures discussed above, this paper introduces a diversity index labeled entropy weighted variance that incorporates correlation insight from a solution set's covariance matrix. The project pairwise selection covariance matrix, equation (66), is a summary of how each selectable element of a dataset varies from the portion of times it was selected (i.e., mean) in relation to the other selectable elements of the dataset. If two selectable elements vary from their mean in opposing (similar) directions, then the covariance between the two selectable elements is negative (positive). Likewise, if two selectable elements vary from the mean randomly or out of sync with respect to one another the covariance tends to zero. The covariance matrix is symmetric and positive semi-definite (Horn and Johnson 1985).

$$\operatorname{cov}(S_{n\times c}) = \begin{bmatrix} s_{1,1}^2 & s_{1,2}^2 & \cdots & s_{1,c}^2 \\ s_{2,1}^2 & s_{2,2}^2 & \cdots & s_{2,c}^2 \\ \vdots & \vdots & \ddots & \vdots \\ s_{c,1}^2 & s_{2,c}^2 & \cdots & s_{c,c}^2 \end{bmatrix} \tag{66}$$

$$s_{i,j}^2 = \frac{1}{n-1}\sum_{k=1}^{n}\left(S_{k,i} - \mu_i\right)\left(S_{k,j} - \mu_j\right) \tag{67}$$

Principal component analysis uses eigenvectors of the covariance matrix to reduce the number of dimensions in the original dataset (Dillon and Goldstein 1982). Eigenvectors of the covariance matrix are the vectors, $\boldsymbol{v}$, that satisfy equation (39). Each eigenvector corresponds to an eigenvalue, λ. Since the covariance matrix is positive semi-define, the eigenvalues are non-negative. The eigenvectors of the covariance matrix represent a set of orthogonal directions that explain the variance within the original dataset. The direction of the eigenvectors account for how the dataset's selectable elements are included in portfolio solutions together. The eigenvector corresponding to the largest eigenvalue represents the transformed elements that have the strongest correlation in the dataset.

$$\text{cov}(S)\boldsymbol{v} = \lambda\boldsymbol{v} \tag{68}$$

Principal component analysis looks to reduce a dataset dimensionality by rotating a multivariate dataset to a less-dimensional dataset while retaining the majority of the variance. The reduced, orthogonal dimensions (i.e., the components) represent hidden, underlying factors of the variance observed in the dataset. Given that the sum of the covariance matrix eigenvalues equals the sum of variance and that the eigenvalues represent the amount of variance explained by the corresponding eigenvector, the portions of the total variance explained by each underlying component can be determined. For example, if the first two largest eigenvalues sum to a value 80% of the total variance, the transformed components represent or retain 80% of the total variance in the dataset. The number of components (i.e., the number of non-zero eigenvalues) represents an alternative view of the "richness" (R) of a dataset. This fact is used with information entropy concepts to quantify diversity considering the number of components that it takes to explain all the variance and the dispersion of the variance among the underlying components.

Consider a decision maker reviewing a solution set in a DSS. In the act of the reviewing the solution set, the decision maker reviews the solution set for differences being communicated through the DSS. How much information does it take to communicate the source of these differences? The number of differences is quantified by the total number of pairwise differences (note the direct scale relationship to the total variance measure). Considering every pairwise difference unit being linked to an underlying-hidden component, the portion of these units linked to an underlying component are represented by the normalized eigenvalues. Assuming the pairwise difference units are communicated or focused upon randomly and independently, the amount of information needed to communicate the specific source of all the variance units in a given dataset from the underlying component perspective is quantified using information entropy.

Let $\boldsymbol{\lambda}$ represent the vector of eigenvalues of $\mathrm{cov}(S)$. Note that the covariance matrix versus the correlation matrix is used to retain the relationship to the binary decision variable (selecting an element). Let $\boldsymbol{\lambda}^{\mathbf{n}}$ represent the non-zero, normalized vector of eigenvalues (i.e., each eigenvalues is scaled by $1/{}^{v}D(S)$). The total variance is analogous to the expected number of messages needed to explain one of the solution's differences from the solution set mean. The value, denoted as variance entropy, equation (40), is analogous to the expected size of a message distinguishing the underlying source component of a difference. Multiplying both results in the expected amount of information needed to distinguish the sources of the differences from the dataset's mean of a single lineup within a solution set (in bits given $\log_2$). Equation (41) is labeled entropy weighted variance. Table 10 shows the results for the solution set examples presented earlier. The value denoted as the entropy weighted total pairwise differences, equation (71), represents the total amount of information needed to explain all the differences' sources in the dataset. This paper proposes the use of entropy weighted variance to compare the diversity of solutions sets of the same size. The paper proposes the use of entropy weighted total pairwise differences for comparing solution sets of different sizes.

TABLE 10. SOLUTION SET EXAMPLES WITH VARIANCE ENTROPY AND WEIGHTED VARIANCE MEASUREMENTS

S	${}^{\mathrm{v}}D$	**Non-zero Eigenvalues ($\boldsymbol{\lambda}$)**	**Variance Entropy ($H(\boldsymbol{\lambda})$)**	**Entropy weighted variance (${}^{\mathrm{I}}D$)**
[1,1,1,0,0,0,0,0,0,0,0,0], [1,1,0,1,0,0,0,0,0,0,0,0], [1,1,0,0,1,0,0,0,0,0,0,0], [1,1,0,0,0,1,0,0,0,0,0,0]	1	0.333, 0.333, 0.333	1.585	1.585
[1,1,0,0,1,0,0,0,0,0,0,0], [1,1,0,0,0,1,0,0,0,0,0,0], [0,0,1,1,1,0,0,0,0,0,0,0], [0,0,1,1,0,1,0,0,0,0,0,0]	2	1.333, 0.667	0.918	1.837
[1,1,0,0,0,1,0,0,0,0,0,0], [1,0,1,0,1,0,0,0,0,0,0,0], [0,1,0,1,1,0,0,0,0,0,0,0], [0,0,1,1,0,1,0,0,0,0,0,0]	2	0.667, 0.667 , 0.667	1.585	3.169
[1,1,1,0,0,0,0,0,0,0,0,0], [0,0,0,1,1,1,0,0,0,0,0,0], [0,0,0,0,0,0,1,1,1,0,0,0], [0,0,0,0,0,0,0,0,0,1,1,1]	3	1.000, 1.000, 1.000	1.585	4.755

Note the diversity index incorporates both the revised richness (number of components with non-zero eigenvalues), the evenness (entropy of the normalized eigenvalues), and the abundance of differences (the total variance of the solution set). Using just the variance entropy as a measure of diversity ignores the abundance of component differences. This is evident in the presented examples in Table 10.

$$H(\boldsymbol{\lambda}) = -\sum_{i}^{|\boldsymbol{\lambda}|} \lambda_i^{\mathrm{n}} \log_2(\lambda_i^{\mathrm{n}}) \tag{69}$$

$${}^{\mathrm{I}}D(S) = \left(\sum_{i}^{|\boldsymbol{\lambda}|} \lambda_i\right) H(\boldsymbol{\lambda}) = \frac{{}^{\mathrm{PWA}}D(S)}{2} H(\boldsymbol{\lambda}) \tag{70}$$

$$ {}^{\mathrm{TI}}D(S) = 2\binom{n}{2}\left(\sum_{i}^{|\lambda|}\lambda_i\right)H(\boldsymbol{\lambda}) = {}^{\mathrm{PWT}}D(S)\,H(\boldsymbol{\lambda}) \tag{71} $$

The entropy weighted variance possesses a few desirable properties compared to the “pure” diversity indexes. Unlike the Solow-Polasky measure, entropy weighted variance does not require the specification and calibration of an additional function to produce a relatable diversity index. The experimentation datasets did not encounter any computation issues computing the entropy weighted variance unlike the Weitzman dynamic programming-based measure. Note that the entropy weighted variance does retain the limitations in regard to measurement of diversity in continuous based decision spaces as visually presented by (Ulrich, Bader, and Thiele 2010). This limits the applicability in using entropy weighted variance to quantify diversity of a set of things based on non-binary attributes.

8.4 Conclusions

This paper provides a foundation for a diversity index from information theory and shows that, without additional calibration as compared to the Solaw-Polasky measure, the entropy weighted variance measure overcomes the failure of the direct pairwise difference based measures to account for correlation in solution sets diversity measurements.

8.5 Proofs and discussion

Proof that the average number of pair-wise differences (or hamming distance) per binary multi-variate dataset equals $2\,{}^{\mathrm{V}}D(x)$.

First, it is shown this relationship applies for a single dimension (column) of the multi-variate dataset. Let $n_1(n_0)$ represent the count of ones (zeros) in the given column and μ represent the average value. Note that when $n_1 = 0$ or $n_0 = 0$ both the variance and pair-wise difference are 0 for any n. To consider the cases when $n_1 > 0$ and $n_0 > 0$, the equations are shown to be equal (see below).

$$\frac{n_1 n_0}{\binom{n}{2}} = 2\,\frac{1}{n-1}\left(\sum_i^n (x_i - \mu)^2\right)$$

$$\frac{n_1(n-n_1)}{\binom{n}{2}} = 2\,\frac{1}{n-1}\left(\sum_i^n \left(x_i - \left(\frac{n_1}{n}\right)\right)^2\right)$$

$$\frac{n_1(n-n_1)}{\binom{n}{2}} = 2\,\frac{1}{n-1}\left(n_1\left(1-\left(\frac{n_1}{n}\right)\right)^2 + (n-n_1)\left(\frac{n_1}{n}\right)^2\right)$$

$$\frac{n_1(n-n_1)}{\binom{n}{2}} = 2\,\frac{1}{n-1}\left(n_1 - \frac{2n_1^2}{n} + \frac{n_1^3}{n^2} + \frac{n_1^2}{n} - \frac{n_1^3}{n^2}\right)$$

$$\frac{n_1(n-n_1)}{\binom{n}{2}} = 2\,\frac{1}{n-1}\left(n_1 - \frac{n_1^2}{n}\right)$$

$$\frac{(n-1)}{\binom{n}{2}} = 2\,\frac{1}{n-n_1}\left(1 - \frac{n_1}{n}\right)$$

$$\frac{n(n-1)}{2\binom{n}{2}} = \frac{n-n_1}{n-n_1}$$

$$\frac{n(n-1)}{2\binom{n}{2}} = 1$$

$$\frac{n(n-1)2!\,(n-2)!}{2n!} = 1$$

$$\frac{n(n-1)(n-2)!}{n!} = 1$$

$$\frac{n!}{n!} = 1$$

To complete the proof for a multi-variate dataset, note this applies for each dimension; the sum of equal values results in total values that are equal.

Upper entropy weighted variance bound for a fixed number of selectable elements portfolio problem

Let *n* represent the number of portfolios (rows) in the dataset. Let *k* be a fixed, exact number of elements required to be in each portfolio. Then the maximum diversity index value is

$$k\left(-\sum_{1}^{n-1}\left(\frac{1}{n-1}\log_2\frac{1}{n-1}\right)\right)$$

The proof above showed that the total variance was equal to 0.5 of the average pair-wise difference. Thus, maximizing the average pair-wise difference then maximizes the total variance. Clearly, the maximum average pair-wise difference is $2k$, (no elements overlap from any portfolio in the set). This results in an maximize variance of k. Next, note that the maximum number of components possible in a multi-variate dataset is the number of portfolios in the dataset minus 1. Thus, $n-1$ components that evenly explain variance the total variance provides the maximum variance entropy.

Lower entropy weighted variance bound discussion for a fixed number of selectable elements portfolio problem

An obvious lower bound for the entropy weighted variance for all multi-variate binary datasets is zero. The total variance is always greater than or equal to 0; the variance entropy is also greater than or equal to 0. If all the variance is explainable by 1 underlying component, notice that the variance entropy is zero. Note by restricting the alternative generation technique, duplicate portfolios are not permitted in a solution set.

To minimize the total variance, the average number of pair-wise differences can be minimized. Let n represent the number of portfolios (rows) in the dataset. The minimum number of average pair-wise differences for a fixed size portfolio is 2. Thus, the minimum variance is 1 given the variance relationship proved above. The minimum number of pair-wise differences for an unconstrained size portfolio is 1 and the minimum variance is thus 0.5. To distinguish between n different portfolio from the decision space perspective, at least celing($\log_2(n)$) components (i.e., number of non-zero eigenvalues) are required. This is the minimum richness, R, of a n sized prescription set.

VIV. Bibliography

Abbas, Ali E. 2003. "An Entropy Approach for Utility Assignment in Decision Analysis." In *AIP Conference Proceedings*, AIP, 328–38. http://aip.scitation.org/doi/abs/10.1063/1.1570550.

Abbas, Ali E., and Ronald A. Howard. 2005. "Attribute Dominance Utility." *Decision Analysis* 2(4): 185–206.

Ahn, Byeong Seok. 2011. "Compatible Weighting Method with Rank Order Centroid: Maximum Entropy Ordered Weighted Averaging Approach." *European Journal of Operational Research* 212(3): 552–59. http://dx.doi.org/10.1016/j.ejor.2011.02.017.

Amiri, Morteza Pakdin. 2010. "Project Selection for Oil-Fields Development by Using the AHP and Fuzzy TOPSIS Methods." *Expert Systems with Applications* 37(9): 6218–24. http://linkinghub.elsevier.com/retrieve/pii/S0957417410001429.

Amorim, P., H. O. Günther, and B. Almada-Lobo. 2012. "Multi-Objective Integrated Production and Distribution Planning of Perishable Products." *International Journal of Production Economics* 138(1): 89–101.

Archer, NP, and F Ghasemzadeh. 1999. "An Integrated Framework for Project Portfolio Selection." *International Journal of Project Management* 17(4): 207–16.

Argyris, Nikolaos, J Figueira, and Alec Morton. 2011. "Interactive Multicriteria Methods for Portfolio Decision Analysis." In *Portfolio Decision Analysis: Improved Methods for Resource Allocation*, eds. Ahti Salo, Jeffrey Keisler, and Alec Morton. New York, NY: Springer, 107–30.

Askin, Ronald G. 2003. "Multiproject Selection and Scheduling with Capacitated , Renewable Resources." In *IIE Annual Conference*, Institute of Industrial and Systems Engineers, 1–6.

Baker, Norman, and James Freeland. 1975. "Recent Advances in R&D Benefit Measurement and Project Selection Methods." *Management Science* 21(10): 1164–75.

Band, R. A. La, and J. D. Andrews. 2004. "Phased Mission Modelling Using Fault Tree Analysis." *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering* 218(2): 83–91.

Basu Roy, Senjuti et al. 2015. "Task Assignment Optimization in Knowledge-Intensive Crowdsourcing." *VLDB Journal* 24(4): 467–91. http://dx.doi.org/10.1007/s00778-015-0385-2.

Baugh Jr., J.W., S.C. Caldwell, and E.D. Brill Jr. 1997. "A Mathematical Programming Approach for Generating Alternatives in Discrete Structural Optimization." *Engineering Optimization* 28(1–2): 1–31.

Beaujon, George, Samuel Marin, and Gary McDonald. 2001. "Balancing and Optimizing a Portfolio of R & D Projects." *Naval Research Logistics* 48(1): 18–40.

Berntsen, Philip B., and Evelina Trutnevyte. 2017. "Ensuring Diversity of National

Energy Scenarios: Bottom-up Energy System Model with Modeling to Generate Alternatives." *Energy* 126: 886–98. http://dx.doi.org/10.1016/j.energy.2017.03.043.

Bessai, Kahina, and François Charoy. 2017. "Business Process Tasks-Assignment and Resource Allocation in Crowdsourcing Context." In *Proceedings - 2016 IEEE 2nd International Conference on Collaboration and Internet Computing, IEEE CIC 2016*, IEEE, 11–18.

Beynon, Malcolm, and Bruce Curry. 2000. "The Dempster-Shafer Theory of Evidence: An Alternative Approach to Multicriteria Decision Modelling." *Omega* 28(February): 37–50. http://www.sciencedirect.com/science/article/pii/S030504839900033X.

Bhattacharyya, Rupak, Pankaj Kumar, and Samarjit Kar. 2011. "Fuzzy R&D Portfolio Selection of Interdependent Projects." *Computers and Mathematics with Applications* 62(10): 3857–70. http://dx.doi.org/10.1016/j.camwa.2011.09.036.

Bickel, J Eric, and James E Smith. 2006. "Optimal Sequential Exploration: A Binary Learning Model." *Decision Analysis* 3(1): 16–32.

Bilbao, A., M. Arenas, M. V. Rodríguez, and J. Antomil. 2007. "On Constructing Expert Betas for Single-Index Model." *European Journal of Operational Research* 183(2): 827–47.

Blecic, Ivan, Arnaldo Cecchini, and Giuseppe A. Trunfio. 2012. "Selection and Scheduling Problem in Continuous Time with Pairwise-Interdependencies." In *International Conference on Computational Science and Its Applications*, Springer, Berlin, Heidelberg, 481–91.

Bobbio, A., L. Portinale, M. Minichino, and E. Ciancamerla. 2001. "Improving the Analysis of Dependable Systems by Mapping Fault Trees into Bayesian Networks." *Reliability Engineering and System Safety* 71(3): 249–60.

Bonini, C. P. 1975. "Risk Evaluation of Investment Projects." *Omega, The Int. Jl of Mgmt Sci.* 3(6): 735–50.

Brill, E.D., J.M. Flach, L.D. Hopkins, and S. Ranjithan. 1990. "MGA: A Decision Support System for Complex, Incompletely Defined Problems." *IEEE Transactions on Systems, Man, and Cybernetics* 20(4): 745–57. http://ieeexplore.ieee.org/document/105076/.

Brill, E Downey, Shoou-Yuh Chang, and Lewis D Hopkins. 1982. "Modeling to Generate Alternatives: The HSJ Approach and an Illustration Using a Problem in Land Use Planning." *Management Science* 28(3): 221–35. http://dx.doi.org/10.1287/mnsc.28.3.221.

Brooks, Fredrick P. Jr. 1995. *The Mythical Man-Month*. Boston, MA: Addison-Wesley Longman Publishing Co., Inc.

Brown, David B, and James E Smith. 2013. "Optimal Sequential Exploration: Bandits, Clairvoyants, and Wildcats." *Operations Research* 61(3): 644–65. http://pubsonline.informs.org/doi/abs/10.1287/opre.2013.1164.

Brown, Gerald G., Robert F. Dell, and Alexandra M. Newman. 2004. "Optimizing Military Capital Planning." *Interfaces* 34(6): 415–25.

Brucker, Peter et al. 1999. "Resource-Constrained Project Scheduling: Notation, Classification, Models, and Methods." *European Journal of Operational Research* 112(1): 3–41.

Bullock, Richard K. 2006. "Theory of Effectiveness Measure (PhD Dissertation)." Air Force Insititue of Technology. http://www.dtic.mil/get-tr-doc/pdf?AD=ADA456717.

Carazo, Ana F. et al. 2010. "Solving a Comprehensive Model for Multiobjective Project Portfolio Selection." *Computers and Operations Research* 37(4): 630–39.

Čepin, Marko, and Borut Mavko. 2002. "A Dynamic Fault Tree." *Reliability Engineering & System Safety* 75(1): 83–91. http://www.sciencedirect.com/science/article/pii/S0951832001001211.

Chang, Da-Yong. 1996. "Applications of the Extent Analysis Method on Fuzzy AHP." *European Journal of Operational Research* 95(3): 649–55.

Chang, Shoou-Yuh, E Downey Brill, and Lewis D. Hopkins. 1982. "Use of Mathematical Models to Generate Alternative Solutions to Water Resources Planning Problems." *Water Resources Research* 18(1): 58–64.

Charnes, A., W. W. Cooper, and E. Rhodes. 1978. "Measuring the Efficiency of Decision Making Units." *European Journal of Operational Research* 2(6): 429–44.

Chen, Jiaqiong. 2005. "Project Selection, Scheduling and Resource Allocation for Engineering Design Groups (PhD Dissertation)." The University of Arizona. https://repository.arizona.edu/bitstream/handle/10150/195457/azu_etd_1157_sip1_m.pdf?sequence=1.

Chen, Jiaqiong, and Ronald G. Askin. 2009. "Project Selection, Scheduling and Resource Allocation with Time Dependent Returns." *European Journal of Operational Research* 193(1): 23–34.

Chen, WN, and Jun Zhang. 2013. "Ant Colony Optimization for Software Project Scheduling and Staffing with an Event-Based Scheduler." *Software Engineering, IEEE Transactions …* 39(1): 1–18. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6165315.

Cochran, Jeffery K., Shwu Min Horng, and John W. Fowler. 2003. "A Multi-Population Genetic Algorithm to Solve Multi-Objective Scheduling Problems for Parallel Machines." *Computers and Operations Research* 30(7): 1087–1102.

Coffin, Mark A., and Bernard W. Taylor. 1996. "Multiple Criteria R&D Project Selection and Scheduling Using Fuzzy Logic." *Computers & Operations Research* 23(3): 207–20.

Crama, Yves, and Michaël Schyns. 2001. "Simulated Annealing for Complex Portfolio Selection Problems." *European Journal of Operational Research* 7(3): 546–571. http://www.sciencedirect.com/science/article/pii/S0377221702007841%5Cnhttp://w ww.rogp.hec.ulg.ac.be/Crama/Publications/Papers/SAPortfo.pdf.

Danielson, Mats, Love Ekenberg, Jim Idefeldt, and Aron Larsson. 2007. "Using a Software Tool for Public Decision Analysis: The Case of Nacka Municipality." *Decision Analysis* 4(2): 76–90. http://pubsonline.informs.org/doi/abs/10.1287/deca.1070.0088.

Danna, Emilie, and David L. Woodruff. 2009. "How to Select a Small Set of Diverse Solutions to Mixed Integer Programming Problems." *Operations Research Letters* 37(4): 255–60. http://dx.doi.org/10.1016/j.orl.2009.03.004.

Debels, Dieter, and Mario Vanhoucke. 2007. "A Decomposition-Based Genetic Algorithm for the Resource-Constrained Project-Scheduling Problem." *Operations Research* 55(3): 457–69. http://pubsonline.informs.org/doi/abs/10.1287/opre.1060.0358.

DeCarolis, J. F., S. Babaee, B. Li, and S. Kanungo. 2016. "Modelling to Generate Alternatives with an Energy System Optimization Model." *Environmental Modelling and Software* 79: 300–310.

Dickinson, Michael W., Anna C. Thornton, and Stephen Graves. 2001. "Technology Portfolio Management: Optimizing Interdependent Projects over Multiple Time Periods." *IEEE Transactions on Engineering Management* 48(4): 518–27.

Dillon, William R., and Matthew Goldstein. 1982. *Multivarate Analysis: Methods and Applications*. New York, NY: Wiley.

Distefano, Salvatore, and Antonio Puliafito. 2009. "Dependability Evaluation with Dynamic Reliability Block Diagrams and Dynamic Fault Trees." *IEEE Transactions on Dependable and Secure Computing* 6(1): 4–17.

Doemer, K.F. et al. 2004. "Pareto Ant Colony Optimization with ILP Processing in Multiobjective Project Portfolio Selection." *European Journal of Operational Research* 171(3): 830–41.

Drineas, Petros et al. 2004. "Clustering Large Graphs via the Singular Value Decomposition." *Machine Learning* 56(1–3): 9–33.

Durga Rao, K. et al. 2009. "Dynamic Fault Tree Analysis Using Monte Carlo Simulation in Probabilistic Safety Assessment." *Reliability Engineering and System Safety* 94(4): 872–83.

Dyer, James S. 1990. "Remarks on the Analytic Hierarchy Process." *Management Science* 36(3): 249–58. http://pubsonline.informs.org/doi/abs/10.1287/mnsc.36.3.249.

Dyer, James S, and Rakesh K. Sarin. 1979. "Measurable Multiattribute Value Functions." *Operations Research* 27(4): 810–22. http://pubsonline.informs.org/doi/abs/10.1287/opre.27.4.810.

Eilat, Harel, Boaz Golany, and Avraham Shtub. 2006. "Constructing and Evaluating Balanced Portfolios of R&D Projects with Interactions: A DEA Based Methodology." *European Journal of Operational Research* 172(3): 1018–39.

Erkut, E., C. ReVelle, and Y. Ülküsal. 1996. "Integer-Friendly Formulations for the r-Separation Problem." *European Journal of Operational Research* 92(2): 342–51.

Ewing, Paul L, William Tarantino, and Gregory Parnell. 2006. "Use of Decsion Analysis in the Army Base Realignment and Closure (BRAC) 2005 Military Value Analysis." *Decision Analysis* 3(1): 33–49.

Fliedner, Thomas, and Juuso Liesiö. 2016. "Adjustable Robustness for Multi-Attribute Project Portfolio Selection." *European Journal of Operational Research* 252(3): 931–46.

Fox, G Edward, Norman R Baker, and John L Bryant. 1984. "Economic Models for R and D Project Selection in the Presence of Project Interactions." *Management Science* 30(7): 890–902.

Ghorbani, S., and M. Rabbani. 2009. "A New Multi-Objective Algorithm for a Project Selection Problem." *Advances in Engineering Software* 40(1): 9–14.

Golabi, Kamal, Craig W Kirkwood, and Alan Sicherman. 1981. "Selecting a Portfolio of Solar Energy Projects Using Multiattribute Preference Theory." *Management Science* 27(2): 174–89. http://pubsonline.informs.org/doi/abs/10.1287/mnsc.27.2.174.

Gomory, Ralph. 1958. "Outline of an Algorithm for Integer Solutions to Linear Programs." *Bulletin of the American Mathematical society* 64(5): 275–78. http://pages.stern.nyu.edu/~rgomory/academic_papers/07_outline.pdf.

Graves, Samuel B., Jeffrey L. Ringuest, and Jonathan F. Bard. 1992. "Recent Developments in Screening Methods for Nondominated Solutions in Multiobjective Optimization." *Computers and Operations Research* 19(7): 683–94.

Gren, Lucas. 2017. "A Fourth Explanation to Brooks' Law - the Aspect of Group Developmental Psychology." In *Proceedings of the 10th International Workshop on Cooperative and Human Aspects of Software Engineering*, IEEE Press, 86–89.

Gröwe-Kuska, Nicole, Holger Heitsch, and Werner Römisch. 2003. "Scenario Reduction and Scenario Tree Construction for Power Management Problems." In *Power Tech Conference Proceedings, 2003 IEEE Bologna*, IEEE, 152–58.

Guikema, Seth D, and Mark W Milke. 2003. "Sensitivity Analysis for Multi-Attribute Project Selection Problems." *Civil Engineering & Environmental Systems* 20(3): 143–62.

Gustafsson, Janne, and Ahti Salo. 2005. "Contingent Portfolio Programming for the Management of Risky Projects." *Operations Research* 53(6): 946–56.

Gutjahr, Walter J. et al. 2010. "Multi-Objective Decision Analysis for Competence-Oriented Project Portfolio Selection." *European Journal of Operational Research* 205(3): 670–79. http://dx.doi.org/10.1016/j.ejor.2010.01.041.

Hartmann, Sönke, and Dirk Briskorn. 2010. "A Survey of Variants and Extensions of the Resource-Constrained Project Scheduling Problem." *European Journal of Operational Research* 207(1): 1–14. http://dx.doi.org/10.1016/j.ejor.2009.11.005.

Hassanzadeh, Farhad, Mohammad Modarres, Hamid R. Nemati, and Kwasi Amoako-Gyampah. 2014. "A Robust R&D Project Portfolio Optimization Model for Pharmaceutical Contract Research Organizations." *International Journal of Production Economics* 158: 18–27. http://dx.doi.org/10.1016/j.ijpe.2014.07.001.

Heidenberger, Kurt, and Christian Stummer. 1999. "Research and Development Project Selection and Resource Allocation: A Review of Quantitative Modelling Approaches." *International Journal of Management Reviews* 1(2): 197–224. http://doi.wiley.com/10.1111/1468-2370.00012.

Hennen, Maike, Matthias Lampe, Philip Voll, and André Bardow. 2017. "SPREAD – Exploring the Decision Space in Energy Systems Synthesis." *Computers and Chemical Engineering* 106: 297–308. http://dx.doi.org/10.1016/j.compchemeng.2017.06.002.

Henriksen, Anne Depiante, and Ann Jensen Traynor. 1999. "A Practical R&D Project-Selection Scoring Tool." *IEEE Transactions on Engineering Management* 46(2): 158–70.

Hess, S W. 1962. "A Dynamic Programming Approach to R and D Budgeting and Project Selection." *IRE Transactions on Engineering Management* (4): 170–79.

Hill, M O. 1973. "Diversity and Evenness: A Unifying Notation and Its Consequences." *Ecology* 54(2): 427–432.

Hill, Raymond R., Alex J. Gutman, Stephen P. Chambal, and Jerry W. Kitchen. 2013. "Acquisition and Testing, DT/OT Testing: The Need for Two-Parameter Requirements." *Quality and Reliability Engineering International* 29(5): 691–97.

Horn, Roger A., and Charles A. Johnson. 1985. *Matrix Analysis*. New York, NY: Cambridge University Press.

Hsia, Pei, Chih-tung Hsu, and David C Kung. 1999. "Brooks' Law Revisited: A System Dynamics Approach." In *Computer Software and Applications Conference, 1999. COMPSAC'99. Proceedings. The Twenty-Third Annual International*, IEEE, 370–75.

Hu, G., L. Wang, S. Fetch, and B. Bidanda. 2008. "A Multi-Objective Model for Project Portfolio Selection to Implement Lean and Six Sigma Concepts." *International Journal of Production Research* 46(23): 6611–25.

Huang, Xiaoxia. 2007. "Optimal Project Selection with Random Fuzzy Parameters." *International Journal of Production Economics* 106(2): 513–22.

———. 2008. "Portfolio Selection with a New Definition of Risk." *European Journal of Operational Research* 186(1): 351–57.

Huynh, Candice H., and Jay Simon. 2016. "Using Means Objectives to Present Risk Information." *Decision Analysis* 13(2): 117–27.

Jiang, Jiang et al. 2011. "Weapon System Capability Assessment under Uncertainty Based on the Evidential Reasoning Approach." *Expert Systems with Applications* 38(11): 13773–84. http://dx.doi.org/10.1016/j.eswa.2011.04.179.

Kangaspunta, Jussi, Juuso Liesiö, and Ahti Salo. 2012. “Cost-Efficiency Analysis of Weapon System Portfolios.” *European Journal of Operational Research* 223(1): 264–75. http://dx.doi.org/10.1016/j.ejor.2012.05.042.

Keeney, Ralph L. 2002. “Common Mistakes in Making Value Trade-Offs.” *Operations Research* 50(6): 935–45.

Keeney, Ralph L., and Howard Raiffa. 1976. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. New York, NY: Wiley.

Keeney, Ralph L. 1981. “Analysis of Preference Dependencies among Objectives.” *Operations Research* 29(6): 1105–20. http://dx.doi.org/10.1287/opre.29.6.1105.

Keisler, Jeffrey. 2004. “Value of Information in Portfolio Decision Analysis.” *Decision Analysis* 1(3): 177–89.

Khorramshahgol, Reza, and Yvon Gousty. 1986. “Delphic Goal Programming (Dgp): A Multi-Objective Cost/Benefit Approach To R&D Portfolio Analysis.” *IEEE transactions on engineering management* 33(3): 172–75.

Kleinmuntz, D. N. 2007. “Resource Allocation Decisions.” In *Advances in Decision Analysis*, eds. Ward Edwards, Ralph F. JrD Miles, and Detlof von Vindterfeldt. Cambridge University Press, 400–418. www.cambridge.org/9780521863681.

Kogan, Konstantin, and Avraham Shtub. 1999. “Scheduling Projects with Variable-Intensity Activities: The Case of Dynamic Earliness and Tardiness Costs.” *European Journal of Operational Research* 118(1): 65–80.

Kolisch, Rainer, and Konrad Meyer. 2006. “Selection and Scheduling of Pharmaceutical Research Projects.” In *Perspectives in Modern Project Scheduling*, Boston, MA: Springer, 321–44. http://link.springer.com/chapter/10.1007/978-0-387-33768-5_13#page-1.

Kopanos, Georgios M., Thomas S. Kyriakidis, and Michael C. Georgiadis. 2014. “New Continuous-Time and Discrete-Time Mathematical Formulation for Resource-Constrained Project Scheduling Problems.” *Computers and Chemical Engineering* 68: 96–106. http://dx.doi.org/10.1016/j.compchemeng.2014.05.009.

Krantz, D.H., R.D. Luce, P. Suppes, and A Tversky. 1971. *Foundations of Measurement, Vol. I: Additive and Polynomial Representations*. New York, NY: Academic Press.

Kuei H. 1994. “A Strategic Decision Model for the Selection of Advance Technology.” *International Journal of Production Research* 32(9): 2117–30.

Lawrence, David B. 2012. *The Economic Value of Information*. New York, NY: Springer New York.

Leachman, Robert C., Abdurrezak Dtncerler, and Sooyoung Kim. 1990. “Resource-Constrained Scheduling of Projects with Variable-Intensity Activities.” *IIE Transactions* 22(1): 31–40.

Lee, Jin Woo, and Soung Hie Kim. 2001. “An Integrated Approach for Interdependent Information System Project Selection.” *International Journal of Project*

Management 19(2): 111–18.

Li, Feng et al. 2012. “Parameterized Model and Approach for Constrained Project Portfolio Optimization.” In *Proceedings of 2012 IEEE International Conference on Service Operations and Logistics, and Informatics, SOLI 2012,* , 462–67.

Li, Xingmei et al. 2015. “An Extended Model for Project Portfolio Selection with Project Divisibility and Interdependency.” *Journal of Systems Science and Systems Engineering* 25(1): 119–38. http://link.springer.com/10.1007/s11518-015-5281-1.

Liang, Chao, and Qing Li. 2008. “Enterprise Information System Project Selection with Regard to BOCR.” *International Journal of Project Management* 26(8): 810–20. http://dx.doi.org/10.1016/j.ijproman.2007.11.001.

Liesiö, Juuso. 2014. “Measurable Multiattribute Value Functions for Portfolio Decision Analysis.” *Decision Analysis* 11(1): 1–20.

Liesiö, Juuso, Pekka Mild, and Ahti Salo. 2007. “Preference Programming for Robust Portfolio Modeling and Project Selection.” *European Journal of Operational Research* 181(3): 1488–1505.

———. 2008. “Robust Portfolio Modeling with Incomplete Cost Information and Project Interdependencies.” *European Journal of Operational Research* 190(3): 679–95.

Liesiö, Juuso, and Antti Punkka. 2014. “Baseline Value Specification and Sensitivity Analysis in Multiattribute Project Portfolio Selection.” *European Journal of Operational Research* 237(3): 946–56.

Liesiö, Juuso, and Ahti Salo. 2012. “Scenario-Based Portfolio Selection of Investment Projects with Incomplete Probability and Utility Information.” *European Journal of Operational Research* 217(1): 162–72.

Liu, Shu Shun, and Chang Jung Wang. 2011. “Optimizing Project Selection and Scheduling Problems with Time-Dependent Resource Constraints.” *Automation in Construction* 20(8): 1110–19. http://dx.doi.org/10.1016/j.autcon.2011.04.012.

Loch, C H, and S Kavadias. 2002. “Dynamic Portfolio Selection of NPD Programs Using Marginal Returns.” *Management Science* 48(10): 1227–41.

Loughlin, Daniel H., S. Ranji Ranjithan, E Downey Brill, and John W. Baugh. 2001. “Genetic Algorithm Approaches for Addressing Unmodeled Objectives in Optimization Problems.” *Engineering Optimization* 33(5): 549–69. http://www.tandfonline.com/doi/abs/10.1080/03052150108940933.

Lucas, Rico et al. 2017. “A Critical Evaluation of Ecological Indices for the Comparative Analysis of Microbial Communities Based on Molecular Datasets.” *FEMS Microbiology Ecology* 93(1). https://academic.oup.com/femsec/article-lookup/doi/10.1093/femsec/fiw209.

Luce, R.Duncan, and John W. Tukey. 1964. “Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement.” *Journal of Mathematical Psychology* 1(1): 1–27.

Maheswari, J. Uma, and Koshy Varghese. 2005. “Project Scheduling Using Dependency Structure Matrix.” *International Journal of Project Management* 23(3): 223–30.

Makowski, D, E. M. T. Hendrix, M. K. van Ittersum, and W. A. H. Rossing. 2000. “A Framework to Study Nearly Optimal Solutions of Linear Programming Models Developed for Agricultural Land Use Exploration.” *Ecological Modelling* 131(1): 65–77. http://linkinghub.elsevier.com/retrieve/pii/S0304380000002490.

Makowski, David, Eligius M T Hendrix, Martin K. Van Ittersum, and Walter A H Rossing. 2001. “Generation and Presentation of Nearly Optimal Solutions for Mixed-Integer Linear Programming, Applied to a Case in Farming System Design.” *European Journal of Operational Research* 132(2): 425–38.

Malakooti, Behnam. 1991. “Measurable Value Functions for Ranking and Selection of Groups of Alternatives.” *Journal of Mathematical Psychology* 35: 92–99.

Mandakovic, Tomislav, and William E Souder. 1985. “An Interactive Decomposable Heuristic for Project Selection.” *Management Science* 31(10): 1257–71.

Mavridis, Panagiotis, David Gross-Amblard, and Zoltán Miklós. 2016. “Using Hierarchical Skills for Optimized Task Assignment in Knowledge-Intensive Crowdsourcing.” In *Proceedings of the 25th International Conference on World Wide Web - WWW ’16*, International World Wide Web Conferences Steering Committee, 843–53. http://dl.acm.org/citation.cfm?doid=2872427.2883070.

Mcdonald, Daniel G., and John Dimmick. 2003. “The Conceptualization and Measurement of Diversity.” *Communication Research* 30(1): 60–79. http://journals.sagepub.com/doi/10.1177/0093650202239026.

Meade, L.M., and A. Presley. 2002. “R&D Project Selection Using the Analytic Network Process.” *IEEE Transactions on Engineering Management* 49(1): 59–66.

Medaglia, Andrés L., Samuel B. Graves, and Jeffrey L. Ringuest. 2007. “A Multiobjective Evolutionary Approach for Linearly Constrained Project Selection under Uncertainty.” *European Journal of Operational Research* 179(3): 869–94.

Medaglia, Andrés L., Darrell Hueth, Juan Carlos Mendieta, and Jorge A. Sefair. 2008. “A Multiobjective Model for the Selection and Timing of Public Enterprise Projects.” *Socio-Economic Planning Sciences* 42(1): 31–45.

Meindl, Bernhard, and Matthias Templ. 2012. “Analysis of Commercial and Free and Open Source Solvers for Linear Optimization Problems.” *ESSnet on commom tools and harmonised methodology for SDC in the ESS* 20. http://www.statistik.tuwien.ac.at/forschung/CS/CS-2012-1complete.pdf.

Montiel, Luis V., and J. Eric Bickel. 2014. “A Generalized Sampling Approach for Multilinear Utility Functions given Partial Preference Information.” *Decision Analysis* 11(3): 147–70.

Morse, J. N. 1980. “Reducing the Size of the Nondominated Set Pruning by Clustering.” *Computers & Operations Research* 7: 55–66.

Naber, Anulark, and Rainer Kolisch. 2014. “MIP Models for Resource-Constrained

Project Scheduling with Flexible Resource Profiles." *European Journal of Operational Research* 239(2): 335–48. http://dx.doi.org/10.1016/j.ejor.2014.05.036.

Nemhauser, G. L., and Z. Ullmann. 1969. "Discrete Dynamic Programming and Capital Allocation." *Management Science* 15(9): 494–505.

Nowak, Maciej. 2013. "Project Portfolio Selection Using Interactive Approach." *Procedia Engineering* 57: 814–22. http://dx.doi.org/10.1016/j.proeng.2013.04.103.

Pahl, Julia, and Stefan Voß. 2014. "Integrating Deterioration and Lifetime Constraints in Production and Supply Chain Planning: A Survey." *European Journal of Operational Research* 238(3): 654–74. http://dx.doi.org/10.1016/j.ejor.2014.01.060.

Pape, Tom. 2017. "Value of Agreement in Decision Analysis: Concept, Measures and Application." *Computers and Operations Research* 80: 82–93.

Petit, Yvan. 2012. "Project Portfolios in Dynamic Environments: Organizing for Uncertainty." *International Journal of Project Management* 30(5): 539–53. http://dx.doi.org/10.1016/j.ijproman.2011.11.007.

Phillips, L., and C. Bana e Costa. 2007. "Transparent Priorisation, Budgeting and Resource Allocation with Multicriteria Decision Analysis and Decision Conferencing." *Annals of operations research* 154(1): 51–68.

Piper, Christopher J., and Andris A. Zoltners. 1976. "Some Easy Postoptimality Analysis for Zero-One Programming." *Management Science* 22(7): 759–65.

Preacher, K J K J, P J P J Curran, and D J J Bauer. 2006. "Computational Tools for Probing Interactions in Multiple Linear Regression, Multilevel Modeling, and Latent Curve Analysis." *Journal of Educational and Behavioral Statistics* 31(4): 437–48.

Rabbani, M., M. Aramoon Bajestani, and G. Baharian Khoshkhou. 2010. "A Multi-Objective Particle Swarm Optimization for Project Selection Problem." *Expert Systems with Applications* 37(1): 315–21. http://dx.doi.org/10.1016/j.eswa.2009.05.056.

Rahman, Habibur et al. 2015. "Task Assignment Optimization in Collaborative Crowdsourcing." In *2015 IEEE International Conference on Data Mining (ICDM)*, IEEE, 949–54.

Razmi, Jafar, and Hamed Rafiei. 2010. "An Integrated Analytic Network Process with Mixed-Integer Non-Linear Programming to Supplier Selection and Order Allocation." *International Journal of Advanced Manufacturing Technology* 49(9–12): 1195–1208.

Ringuest, Jeffrey L., and Samuel B. Graves. 1990. "The Linear R&D Project Selection Problem:An Alternative to Net Present Value." *IEEE Transactions on Engineering Management* 37(2): 143–46.

Ringuest, Jeffrey L., Samuel B. Graves, and Randy H. Case. 2004. "Mean-Gini Analysis in R&D Portfolio Selection." *European Journal of Operational Research* 154(1): 157–69.

Rong, Aiying, Renzo Akkerman, and Martin Grunow. 2011. “An Optimization Approach for Managing Fresh Food Quality throughout the Supply Chain.” *International Journal of Production Economics* 131(1): 421–29. http://dx.doi.org/10.1016/j.ijpe.2009.11.026.

Salo, Ahti A., and Raimo P. Hämäläinen. 2001. “Preference Ratios in Multiattribute Evaluation (PRIME)-Elicitation and Decision Procedures under Incomplete Information.” *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans* 31(6): 533–45.

Salo, Ahti, Jeffrey Keisler, and Alec Morton. 2011. *Portfolio Decision Analysis: Improved Methods for Resource Allocation*. New York, NY: Springer.

Salo, Ahti, and Antti Punkka. 2005. “Rank Inclusion in Criteria Hierarchies.” *European Journal of Operational Research* 163(2): 338–56.

Saltelli, Andrea. 2002. “Sensitivity Analysis for Importance Assessment.” *Risk Analysis* 22(3): 579–90.

Santhanam, Radhika, and George J. Kyparisis. 1996. “A Decision Model for Interdependent Information System Project Selection.” *European Journal of Operational Research* 89(2): 380–99.

Sarin, Rakesh Kumar. 1982. “Strength of Preference and Risky Choice.” *Operations Research* 30(5): 982–97.

Sayın, Serpil. 2000. “Measuring the Quality of Discrete Representations of Efficient Sets in Multiple Objective Mathematical Programming.” *Mathematical Programming* 87(3): 543.

Scott, D A, C Reeves, E M Russell, and M Napper. 2009. “Eliciting Public Preferences for Healthcare: A Systematic Review of Techniques.” *Chinese Journal of Evidence-Based Medicine* 9(2): 129–30.

Scott, Dana, and Patrick Suppes. 1958. “Foundational Aspects of Theories of Measurment.” *The journal of symbolic logic* 23(2): 113–28.

Sefair, Jorge A., and Andrés L. Medaglia. 2005. “Towards a Model for Selection and Scheduling of Risky Projects.” In *Proceedings of the 2005 Systems and Information Engineering Design Symposium*, IEEE, 158–64.

Shannon, Claude Edmund. 1948. “A Mathematical Theory of Communication.” *Bell System Technical Journal* 5(3): 3.

Shariatmadari, Mohammad, Nasim Nahavandi, Seyed Hessameddin Zegordi, and Mohammad Hossein Sobhiyah. 2017. “Integrated Resource Management for Simultaneous Project Selection and Scheduling.” *Computers and Industrial Engineering* 109: 39–47. http://dx.doi.org/10.1016/j.cie.2017.04.003.

Shir, Ofer M., Mike Preuss, Boris Naujoks, and Michael Emmerich. 2010. “Enhancing Decision Space Diversity in Evolutionary Multiobjective Algorithms.” In *International Conference on Evolutionary Multi-Criterion Optimization*, Springer, 95–109.

da Silva, Celmar G. et al. 2017. “An Improved Visualization-Based Approach for Project Portfolio Selection.” *Computers in Human Behavior* 73: 685–96.

Sinha, Koushik, Pratham Majumder, and Geetha Manjunath. 2016. “A Dynamic Microtask Scheduling Approach for SLO Based Human-Augmented Computing.” In *2016 3rd International Conference on Recent Advances in Information Technology, RAIT 2016*, IEEE, 191–96.

Smith, Brian, Priya Sharma, and Paula Hooper. 2006. “Decision Making in Online Fantasy Sports Communities.” *Interactive Technology & Smart Education* 3(4): 347–60.

Sobol, I. M., and S. S. Kucherenko. 2005. “Global Sensitivity Indices for Nonlinear Mathematical Models. Review.” *Wilmott* 2005(1): 56–61. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Global+Sensitivity+Indices+for+Nonlinear+Mathematical+Models+.+Review#7.

Solow, Andrew R., and Stephen Polasky. 1994. “Measuring Biological Diversity.” *Environmental and Ecological Statistics* 1(2): 95–103. http://link.springer.com/10.1007/BF02426650.

Spetzler, Carl S., and Carl-Axel S. Staël Von Holstein. 1975. “Exceptional Paper—Probability Encoding in Decision Analysis.” *Management Science* 22(3): 340–58. http://pubsonline.informs.org/doi/abs/10.1287/mnsc.22.3.340.

Srivastava, Joydeep, Terry Connolly, and Lee Roy Beach. 1995. “Do Ranks Suffice? A Comparison of Alternative Weighting Approaches in Value Elicitation.” *Organizational Behavior and Human Decision Processes* 63(1): 112–16. http://linkinghub.elsevier.com/retrieve/pii/S0749597885710667.

States, United, and Government Accountability. 2016. “Additional Steps Could Better Integrate Intelligence Input into DOD’s Acquisition of Major Weapon Systems (Publication No. GAO-17-10).” https://www.gao.gov/products/GAO-17-10.

Steuer, Ralph E., and Frederick W. Harris. 1980. “Intra-Set Point Generation and Filtering in Decision and Criterion Space.” *Computers & Operations Research* 7(1–2): 41–53. http://linkinghub.elsevier.com/retrieve/pii/0305054880900131.

Stummer, Christian, and Kurt Heidenberger. 2003. “Interactive R&D Portfolio Analysis with Project Interdependencies and Time Profiles of Multiple Objectives.” *IEEE Transactions on Engineering Management* 50(2): 175–83. http://ieeexplore.ieee.org/document/1202980/.

Stummer, Christian, Elmar Kiesling, and Walter J. Gutjahr. 2009. “A Multicriteria Decision Support System for Competence-Driven Project Portfolio Selection.” *International Journal of Information Technology & Decision Making* 08(02): 379–401. http://www.worldscientific.com/doi/abs/10.1142/S0219622009003429.

Sun, Hongyi, and Tianchao Ma. 2005. “A Packing-Multiple-Boxes Model for R&D Project Selection and Scheduling.” *Technovation* 25(11): 1355–61. http://linkinghub.elsevier.com/retrieve/pii/S0166497204001166.

Tavana, Madjid, Amir-reza Abtahi, and Kaveh Khalili-damghani. 2014. “A New Multi-Objective Multi-Mode Model for Solving Preemptive Time--Cost--Quality Trade-off Project Scheduling Problems.” *Expert Systems with Applications* 41(4): 1830–46.

Taylor, Bernard W., Laurence J Moore, and Edward R Clayton. 1982. “R&D Project Selection and Manpower Allocation with Integer Nonlinear Goal Programming.” *Management Science* 28(10): 1149–58. http://pubsonline.informs.org/doi/abs/10.1287/mnsc.28.10.1149.

Tofighian, Ali Asghar, and B. Naderi. 2015. “Modeling and Solving the Project Selection and Scheduling.” *Computers & Industrial Engineering* 83: 30–38. http://dx.doi.org/10.1016/j.cie.2015.01.012.

Toppila, Antti, and Ahti Salo. 2017. “Binary Decision Diagrams for Generating and Storing Non-Dominated Project Portfolios with Interval-Valued Project Scores.” *European Journal of Operational Research* 260(1): 244–54. http://linkinghub.elsevier.com/retrieve/pii/S0377221716310530.

Trainor, Tim, Gregory Parnell, and Michael Kwinn. 2004. *USMA Study of the Installation Management Agency CONUS Region Structure (No. DSE-TR-0506).* MILITARY ACADEMY WEST POINT NY DEPT OF SYSTEM ENGINEERING. http://www.dtic.mil/get-tr-doc/pdf?AD=ADA427027.

Tran-Thanh, Long et al. 2014. “BudgetFix : Budget Limited Crowdsourcing for Interdependent Task Allocation with Quality Guarantees.” In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*, , 477–84.

Udell, Madeleine, and Stephen Boyd. 2013. “Maximizing a Sum of Sigmoids.” *Optimization and Engineering*: 1–25. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.699.5817&rep=rep1&type=pdf.

Ulrich, Tamara, Johannes Bader, and Lothar Thiele. 2010. “Defining and Optimizing Indicator-Based Diversity Measures in Multiobjective Search.” In *Parallel Problem Solving from Nature, PPSN XI*, Berlin, Heidelberg: Springer Berlin Heidelberg, 707–17. http://link.springer.com/10.1007/978-3-642-15844-5_71.

Vanhoucke, Mario et al. 2008. “An Evaluation of the Adequacy of Project Network Generators with Systematically Sampled Networks.” *European Journal of Operational Research* 187(2): 511–24. http://linkinghub.elsevier.com/retrieve/pii/S037722170700375X.

Vilkkumaa, Eeva, Ahti Salo, and Juuso Liesiö. 2014. “Multicriteria Portfolio Modeling for the Development of Shared Action Agendas.” *Group Decision and Negotiation* 23(1): 49–70. http://link.springer.com/10.1007/s10726-012-9328-0.

Villarreal, Bernardo, and Mark H. Karwan. 1981. “Multicriteria Integer Programming: A (Hybrid) Dynamic Programming Recursive Approach.” *Mathematical Programming* 21(1): 204–23. http://link.springer.com/10.1007/BF01584241.

Voll, Philip et al. 2015. “The Optimum Is Not Enough: A near-Optimal Solution Paradigm for Energy Systems Synthesis.” *Energy* 82: 446–56. http://linkinghub.elsevier.com/retrieve/pii/S0360544215000791.

Wall, Matthew Bartschi. 1996. MIT “A Genetic Algorithm for Resource-Constrained Scheduling (PhD Dissertation).” Massachusetts Institute of Technology. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.2444&rep=rep1&type=pdf.

Walls, Michael R., G. Thomas Morahan, and James S. Dyer. 1995. “Decision Analysis of Exploration Opportunities in the Onshore US at Phillips Petroleum Company.” *Interfaces* 25(6): 39–56.

Węglarz, Jan. 1981. “Project Scheduling with Continuously-Divisible, Doubly Constrained Resources.” *Management Science* 27(9): 1040–53. http://pubsonline.informs.org/doi/abs/10.1287/mnsc.27.9.1040.

Weitzman, Martin L. 1992. “On Diversity.” *Quarterly Journal of Economics* 107(2): 363–405.

Wilbaut, Christophe, Said Hanafi, and Said Salhi. 2007. “A Survey of Effective Heuristics and Their Application to a Variety of Knapsack Problems.” *IMA Journal of Management Mathematics* 19(3): 227–44. https://academic.oup.com/imaman/article-lookup/doi/10.1093/imaman/dpn004.

Wilck, Joseph et al. 2016. “An Optimization Approach to Balancing Risk and Cost in Combatant Command Capability Advocacy.” *Industrial and Systems Engineering Review* 4(1): 12–21.

Williams, Laurie, Anuja Shukla, and Annie I Antón. 2004. “An Initial Exploration of the Relationship Between Pair Programming and Brooks’ Law.” In *2004 Agile Development Conference*, IEEE, 11–20.

Winterfeldt, D. von, and Ward Edwards. 1986. *Decision Analysis and Behavioral Research*. New York: Cambridge University Press.

You, Byungjun, and Takeo Yamada. 2007. “A Pegging Approach to the Precedence-Constrained Knapsack Problem.” *European Journal of Operational Research* 183(2): 618–32.

Zan, Kun, J Eric Bickel, Kun Zan, and J Eric Bickel. 2013. “Components of Portfolio Value of Information Components of Portfolio Value of Information.” *Decision Analysis* 10(2)(February 2015): 171–85.

Zechman, Emily M, and Ranji S Ranjithan. 2007. “Generating Alternatives Using Evolutionary Algorithms for Water Resources and Environmental Management Problems.” *Journal of Water Resources Planning and Management* 133(April): 156–65.

Zhao, Tianyi, and Xiaoxia Huang. 2017. “Optimizing Project Selection: A Perspective from Resource Constraints.” *Optimization* 66(11): 1863–78. https://www.tandfonline.com/doi/full/10.1080/02331934.2017.1317777.

Zuluaga, Andrea, Jorge A. Sefair, and Andres L. Medaglia. 2007. "Model for the Selection and Scheduling of Interdependent Projects." In *2007 IEEE Systems and Information Engineering Design Symposium*, IEEE, 1–7. https://ieeexplore.ieee.org/document/4374020/.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704–0188*

1. REPORT DATE (DD–MM–YYYY)	2. REPORT TYPE	3. DATES COVERED (*From* — *To*)
13-09-2018	Dissertation	SEP 2009 – SEP 2018

4. TITLE AND SUBTITLE METHODS TO SUPPORT THE PROJECT SELECTION PROBLEM WITH NON-LINEAR PORTFOLIO OBJECTIVES, TIME SENSITIVE OBJECTIVES, TIME SENSITIVE RESOURCE CONSTRAINTS, AND MODELING INADEQUACIES	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Ranly, Neil C., CTR USAF	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB OH 45433-7765	8. PERFORMING ORGANIZATION REPORT NUMBER AFIT-ENS-DS-18-S-040
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Life Cycle Management Center OZA ATTN: Theodore Lewis, PhD 1865 4th Street Wright-Patterson AFB, OH 45433 (937) 904-5654, Theodore.Lewis@us.af.mil	10. SPONSOR/MONITOR'S ACRONYM(S) AFMC LCMC OZA 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Statement A: Approved For Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

The United States Air Force relies upon information production activities to gain insight regarding uncertainties affecting important system configuration and in-mission task execution decisions. Constrained resources that prevent the fulfillment of every information production request, multiple information requestors holding different temporal-sensitive objectives, non-constant marginal value preferences, and information-product aging factors that affect the value-of-information complicate the management of these activities. This dissertation reviews project selection research related to these issues and presents novel methods to address these complications. Quantitative experimentation results demonstrate these methods' significance.

15. SUBJECT TERMS
Project Selection, Resource Allocation, Portfolio Decision Analysis, Diversity, Scheduling

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Dr. Jeff Weir, AFIT/ENS
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include Area Code)
U	U	U	UU	142	(937) 255-6565, ext 4523 (NOT DSN) Jeffery.Weir@aft.edu