Forecasting System Capability Using the Scenario-Based Experimental-Design Test Approach

by Ya Lun Li

B.S. in Electrical Engineering, June 2002, Wright State University
M.S. in Program Management, December 2005, Florida Institute of Technology

A Dissertation submitted to

The Faculty of
The School of Engineering and Applied Science
of The George Washington University
in partial satisfaction of the requirements
of the degree of Doctor of Philosophy

January 10th, 2019

Dissertation directed by

Blake Roberts
Professional Lecturer of Engineering Management & Systems Engineering

Michael Grenn
Professional Lecturer of Engineering Management and Systems Engineering

The School of Engineering and Applied Science of The George Washington University certifies that Ya Lun Li has passed the Final Examination for the degree of Doctor of Philosophy as of September 26th, 2018. This is the final and approved form of the dissertation.

Forecasting System Capability Using the Scenario-Based Experimental-Design Test Approach

Ya Lun Li

Dissertation Research Committee:

Shahram Sarkani, Professor of Engineering Management and Systems Engineering, Committee Member

Michael Grenn, Professional Lecturer of Engineering Management and Systems Engineering, Dissertation Co-Director

Amirhossein Etemadi, Professional Lecture of Engineering Management and Systems Engineering, Committee Member

Thomas Mazzuchi, Professor of Engineering Management and Systems Engineering & Decision Sciences, Committee Member

Blake Roberts, Professional Lecturer of Engineering Management and Systems Engineering, Dissertation Co-Director

Dedications

To my wife, Sarah,

my children Brooke, Jack, Oscar, Aaron, and Audrey,

and my parents

Acknowledgement

I would like to express my sincere appreciation and thanks to my advisors, Dr. Blake Roberts and Dr. Michael Grenn for their invaluable guidance throughout my Ph.D. study at the George Washington University. Their advice and encouragement have been essential to my success.

I also wish to thank the members of my dissertation committee, Drs. Thomas Mazzuchi, Shahram Sarkani, and Amir Etemadi, for their valuable time and constructive suggestions.

I also greatly appreciate the financial support from US Navy's Naval Surface Warfare Center at Dahlgren, VA.

Finally, and most importantly, I want to express my endless gratitude to my parents, Professor LiangYou Li and Mrs. ShouZhen Zhang, and my wife, Sarah, for their love, patience, encouragement, support, and sacrifices. I would also like to thank my children, Brooke, Jack, Oscar, Aaron, and Audrey for the happiness they brought to us in the past three years.

For the motivation and desire that I receive to pursue my dreams and become the best I can, thank you all.

Abstract

The Scenario-Based Experimental-Design Test Approach: Forecasting Higher System-Level Capabilities through Amalgamation of Element Verification Data

During the system-elements' requirement-based test approach (RBT&E), it is difficult for both Systems Engineering (SE) and Test and Evaluation (T&E) communities to objectively identify the corresponding system-of-interest (SOI) capability (or ability to achieve requirements throughout its intended operational environment) through aggregate interpretation and analysis of available element performance and verification data. This is because traditional RBT&E data are collected under unique measurements and verification conditions specific to only the allocated element requirements (e.g., element A's demonstrated y_1-*mph* speed under $x_1{}^oC$ verification condition versus element B's demonstrated payload of y_2-*tons* at x_2-m altitude). Such mismatched variables cannot be quantitatively amalgamated to objectively forecast a comprehensive and representative SOI operational capability, thus reverting analyst and stakeholders to subjectively interpret element verification data and produce imprecise and/or incomplete SOI capability forecasts. In this study, the Scenario-Based Experimental-Design (SBED) test approach has been developed as a supplement to the traditional RBT&E approach to resolve these data-mismatch challenges. Through the implementation of Design of Experiments (DoE) test methods and response surface methodology statistical techniques, SBED models every independent element verification data as respective response models with respect to (w.r.t.) the *"SOI-element capability"* variable (or ability to fulfill corresponding element requirements) all as influenced by the overall SOI's intended operational environment; it then objectively forecasts the SOI capability via either standalone Boolean-algebraic

summation computations of the (now matching) element-performance response models. This study's SBED test approach contributes to both SE and T&E through the introduction of two benefits that have been previously unobtainable during SOI Implementation process. First, the early delivery of objective and quantitative SOI capability forecast(s) to support upfront identification and management of SOI design-related risks and uncertainties. Second, the beneficial delivery of more-descriptive "element capability" response models, which comprehensively characterize the respective element performances throughout the entire SOI operational environment thus expanding the scope of element performance verifications to beyond just allocated requirement-specified conditions. These introduced benefits are demonstrated through a data-analysis example in this dissertation that applies SBED's data-analysis techniques to an existing United States Air Force flight test's data set and compares its SBED-produced quantitative SOI capability forecasts against the test's original qualitative conclusions.

Table of Contents

Dedications ..... **iii**

Acknowledgement ..... **iv**

Abstract ..... **v**

Table of Contents ..... **vii**

List of Illustrations ..... **x**

List of Tables ..... **xi**

List of Abbreviations ..... **xii**

Mathematical Notations ..... **xiii**

1 Introduction to Research 1

1.1 Background 1

1.2 Statement of the Problem 4

1.3 Approach 6

1.4 Organization of the Document 9

1.5 Significance 10

1.6 Scope and Limitations 13

2 Literature Review 16

2.1 The Traditional Roles and Expectations of T&E 16

2.1.1 The Requirement-Based Performance Verification Goal 16

2.1.2 Benefits of The Traditional RBT&E Approach 18

2.1.2.1 Adaptability Towards SOI Design's Hierarchical Complexity 19

2.1.2.2 Suitability to Element-Development Strategy 21

2.1.2.3 Additional SE and T&E Benefits 22

2.1.3 Implementation Process' Inextricability Towards RBT&E 23

2.2 New Roles and Expectations of T&E 24

2.2.1 The SOI-Capability Identification Goal 24

2.2.2 Inadequacies of the RBT&E Approach 25

2.2.3 Inadequacies to Support MBSE's SOI-Capability Simulation 27

2.2.4 Supplementation of the RBT&E Approach 28

2.3 Existing RBT&E Supplements to Forecast SOI Capability 33

2.3.1 SE-Based T&E Framework and Methods 34

2.3.2 The Early Prototype T&E Approach 35
2.3.3 MBSE and M&S Test Approaches 36
2.3.4 Experimental Design Test Method 38
2.3.5 SOI Mission-Relevant Assessment Analysis (Deficiency Tracking) 40
2.3.6 Encapsulation by the Proposed SBED Test Approach 42
3 The SBED Test Approach Methodology 44
3.1 Definitions and Prerequisite Assumptions 45
3.1.1 Defining the 'SOI Capability' 45
3.1.2 SOI Capability-and-Element Capability Hierarchical Relationship 45
3.2 SBED Methodology Description 46
3.2.1 Step 1 Depict the SOI Capability Scenario 46
3.2.2 Step 2 Introduce SOI Operational Environment as the Common Domain 48
3.2.3 Step 3 Design and Execute Element Verifications with DoE Methods 50
3.2.3.1 Initial Screening DoE Test Methods 52
3.2.3.2 Sequential DoE Test Methods 54
3.2.3.3 Alignment to the Primary Element Verification's Goals and Objectives 57
3.2.3.4 Capture Element Performance and Verification Data 59
3.2.4 Step 4 Calculate Standardized *Element-Capability* Forecast Models 60
3.2.4.1 RSM Analysis Setups 60
3.2.4.1.1 Conversion to Coded ζj Input-Variable Values 60
*3.2.4.1.2 Partitioning of **R** for Point-Estimate Calculations* 61
3.2.4.2 Generating Element Performance Mean-Response Models 63
3.2.4.2.1 Calculate Elements' Mean Response 63
3.2.4.2.2 Derive Elements' Prediction Point Estimates 65
3.2.4.2.3 Construct Element Mean-Response Models 67
3.2.4.3 Generate Element Performance Prediction Confidence Intervals 68
3.2.4.3.1 Calculate Elements' Point Estimate Variances 68
3.2.4.3.2 Construct Elements' Prediction-Variability Models 70
3.2.4.4 Infer Element Capability Models 72
3.2.4.4.1 Calculate Element-Capability Point Estimates 73
3.2.4.4.2 Construct Element-Capability Models 76

3.2.4.5 Resolution of RBT&E's Data-Inconsistency Challenges 79
3.2.5 Step 5 Forecasting the SOI-Capability 88
3.2.5.1 Through Superposition of Element-Capability Models 88
3.2.5.2 Inputs to MBSE Models 90
3.3 Benefits of SBED 93
3.3.1 Early Delivery of SOI Capability Forecast Through SBED 93
3.3.2 Enhancement of the Requirement-Based Performance Verification Goal 94
3.3.3 Enablement of MBSE Supportability 94
3.3.4 Preservation of RBT&E's Intrinsic Benefits 95
3.3.5 Improvements to Remaining RBT&E' Intrinsic Benefits 95
3.3.6 Benefits to SE Processes 97
3.4 Limitations of the SBED Test Approach 98
4 Data Analysis Example 100
4.1 AFFTC Test Method Description and Original Conclusions 100
4.2 Overview of the SBED Data Analysis Example 105
4.3 The SBED Data Analysis of the T-38 Test Data 106
4.3.1 Step 1 Develop a Scenario-Based $\mathbf{F}_{OWL-suitabiltiy}$ Definition 106
4.3.2 Step 2 Introduce the Common Domain 107
4.3.3 Step 3 Design and Collect Element Verification Data with DoE 108
4.3.4 Step 4 Calculate *Element-Capability* Forecasts 109
4.3.5 Step 5 Qualitative Amalgamation of SOI-Level Capability Forecast 112
4.4 Validation of the Conclusions 114
5 Discussions, Conclusions, and Future Works 116
5.1 Discussions and Conclusions 116
5.2 Limitations and Future Research Opportunities 117
5.2.1 Limitations and Boundaries of Study 117
5.2.2 Recommendations for Future 117
Bibliography/Reference List 120
Appendix A, Chapter 3's Raw Data Sets 140

List of Illustrations

Figure 1 - 1, The SBED Test Approach's Model Diagram and Data Flow Chart 8
Figure 1 - 2, SBED Information Feedforward- and Feedback Flow within the SE Vee .. 12
Figure 3 - 1, An Arbitrary SOI-Capability Scenario Diagram 48
Figure 3 - 2, A 2-Dimensional SOI Operational Environment Domain 62
Figure 3 - 3, Calculated Mean-Response Point Estimate 66
Figure 3 - 4, A SOI Operational Environment with Mean-Response Point Estimates 68
Figure 3 - 5, Calculated Prediction Confidence-Interval Point Estimate 70
Figure 3 - 6, SOI Operational Environment with Confidence-Interval Point Estimates .. 71
Figure 3 - 7, Derived Element Capability Point Estimate via *t*-Test 75
Figure 3 - 8, Calucated *Element-Capability* Point Estimate 77
Figure 3 - 9, Examples of SBED's Element Response Plots 78
Figure 3 - 10, Contour plot SBED-Producred Element-Capablity Responses 78
Figure 3 - 11, Element f_1's Response Models 81
Figure 3 - 12, Element f_2's Response Models 82
Figure 3 - 13, Element f_3's Response Models 82
Figure 3 - 14, Element f_4's Response Models 83
Figure 3 - 15, Example's four element capability models. 84
Figure 3 - 16, Amalgamation of SOI-Capability via Boolean-Algebraic Summations 92
Figure 4 - 1, The SBED Data Flow Chart as Applied to the Data-Analysis Example 104
Figure 4 - 2, AFFTC Data-Collection (DoE Design Points) Locations 109
Figure 4 - 3, SBED-Produced T-38C *Element-Capability* Models 111
Figure 4 - 4, SBED-Produced T-38 SOI-Level Capability Forecast 113

List of Tables

Table 2 - 1, Examples of Hypothetical Allocated Element Requirements 30
Table 2 - 2, Reporting Formats of the Traditional RBT&E Verification Data.....31
Table 2 - 3, Example of Deficiency Classification Guideline 41
Table 3 - 1, Example of Raw Design of Experiment Performance Verification Data.....59
Table 3 - 2, DoE Design-Point Conversions.....61
Table 3 - 3, DoE Methods' Reported Verification Data.....80
Table 3 - 4, Explanations of *Element-Capability* Response Models 84
Table 3 - 5, Reporting Formats of the Traditional RBT&E Verification Data.....87
Table 4 - 1, AFFTC Test's Standalone Lower-Level Verification Data 101
Table 4 - 2, AFFTC Test's Pilot OWL Survey Ratings..... 102
Table 4 - 3, T38C Test's Observation Region..... 108
Table 4 - 4, T-38C Operational Environment Data-Resolution Definitions..... 108
Table 4 - 5, Linear Regression Model Coefficients for Element Response Models..... 110
Table 4 - 6, USAF FEB 2011T-38C Incident's Ambient Condition 115
Table 4 - 7, Interpolated L_1's OWL-Suitability Point Estimate..... 115

List of Abbreviations

Acronym	Definition
AFFTC	(United States) Air Force Flight Test Center
ANOVA	analysis of variance
CCD	central composite design
CW	cross wind
DoD	Department of Defense
DoE	design of experiments
DoN	Department of Navy
f-wt	fuel weight
GAO	General Accounting Office
HW	head wind
INCOSE	International Council on Systems Engineering
M&S	modeling and simulation
MBSE	model-based Systems Engineering
MODAF	Ministry of Defense
OE	operational effectiveness
OpsCon	operational concept
OS	operational suitability
OWL	operator workload
RBT&E	Requirement-based test & evaluation
RE	Requirements Engineering
RSM	response surface methodology
SBED	the scenario-based experimental-design (test approach)
SE	Systems Engienering
SOI	system of interest
SoS	system of systems
t	temperature
T&E	test and evaluation
USAF	United States Air Force
V&V	verification and validation
w.r.t.	with respect to

Mathematical Notations

x or x^*	scalar values
$\boldsymbol{X}$	matrix notations or arrays
$\widehat{\boldsymbol{F}}_{\boldsymbol{capaiblity}}$	Predicted system-level operational capability
$\boldsymbol{f_i}$	System element-level functionality
$\widehat{\boldsymbol{f}}_{\boldsymbol{i}},\ \widehat{\boldsymbol{y}}_{\boldsymbol{f_i}}$	Predicted regression model of element-level functionality (mean response model)
$\widehat{\boldsymbol{f}}_{\boldsymbol{i_variability}}$	Predicted confidence-interval model of element-level functionality
$\widehat{\boldsymbol{f}}_{\boldsymbol{i_capabiltiy}}$	Predicted capability of element-level functionality
$\boldsymbol{G}$	A set of m elements sequentially ordered to define system capability
k	Number of parameters that define the system's intended operational environment
m	Total number of elements within a set $\boldsymbol{G}$
n	Number of unique conditions within a data set ('design points')
N	Total sample side of data ('trials')
$\boldsymbol{\mathcal{N}}(\ -\ ,\ -\ )$	Normal (Gaussian) distribution
L_x	Designation of T-38C landing technique (part of data-analysis example)
$\boldsymbol{R}$	System's intended operational environment
R_{x_R}	Partitioned subspace of $\boldsymbol{R}$
T_i or T_{f_i}	Element requirement threshold value
$\{T_{f_i_min},\ T_{f_i_max}\}$	Element requirement threshold tolerance
$\boldsymbol{X}$	Design matrix (coded) of a design-of-experiment
x_R	Number of partition intervals, along a single axis (a.k.a. data fidelity)
$\boldsymbol{x_o}$	Moment matrix that identifies a R_{x_R} subspace's respective condition
$\boldsymbol{y_{f_i}}$	Collected output measurements of a $\boldsymbol{f_i}$
ξ_j	Parameter(s) that define $\boldsymbol{R}$ (DoE controlled independent variables or input factors)

1 Introduction to Research

1.1 Background

Throughout the system-of-interest (SOI) development life cycle, the Test and Evaluation (T&E) community's primary goal is to provide objective evidence that the implemented and integrated functions-under-test perform their intended functions and fulfill their allocated performance requirement thresholds and characteristics (International Council on Systems Engineering [INCOSE], 2015; Department of Defense [DoD], 2017). While this *Requirement-Based Performance Verification* T&E goal remains paramount during the SOI Implementation and subsequent Integration-and-Verification processes, there exists another important T&E goal to identify either the SOI's, or the next integrated-level's, capability (i.e., the ability to achieve integrated-level requirements throughout its stated operational environment) as early as possible per INCOSE (2015), DoD (2017), Eigbe, Sauser, & Boardman (2010), and Helle & Schamai (2016). To date, it is difficult for both Systems Engineering (SE) and T&E communities to concurrently accomplish this new *SOI-Capability Identification* goal during the SOI Implementation process and its element verification activities. This is because the requirement-based test approach (RBT&E), that is traditionally employed to verify element performances' compliance with respect to (w.r.t.) allocated element requirements, introduces several data-metrology and -domain inconsistency challenges towards the objective identification of SOI capability.

Specific to accomplishing the primary *Requirement-Based Performance Verification* T&E goal, the employment of the RBT&E test approach objectively confirms that the SOI-element performance satisfies its allocated element requirement thresholds by collecting corresponding performance and verification data specifically at the

requirement's verification conditions (e.g., demonstrated y_1-*mph* speed under $x_1{}^oC$ verification condition, y_2-*tons* payload at x_2-*m* altitude, y_3-*Watts* sustain output power after x_3-*hours* of operation, etc.) according to INCOSE (2015) and reiterated in Bjorkman & Gray (2009). RBT&E's bottom-up reductionist verification approach of generating only test artifacts (e.g., test cases and test scenarios) that are specific to the allocated element requirements and deferring the arduous integrated SOI verification/validation objectives to later (Hesari, Behjati, & Yue, 2013; Bjorkman & Gray, 2009) is concomitant to: (a) the upfront and reliable identification of performance failures and anomalies w.r.t. the allocated element requirement (Helle &Schamai, 2016); (b) the validation of element requirements as written (e.g., quantitative, measurable, unambiguous, and testable) per INCOSE (2015); and (c) the establishment of unambiguous traceability between the two (INCOSE, 2015; Bjorkman & Gray, 2009; Winkler & von Pilgrim, 2009). According to Li, Roberts, & Grenn (2018a) and as further discussed in Subsection 2.1.2 (Benefits of the Traditional RBT&E Approach), these inherent SE and T&E benefits inextricably link the RBT&E test approach (and justifiably so) to the SOI Implementation process and its element-verification activities.

When it comes to concurrently accomplishing the *SOI-Capability Identification* goal during the SOI Implementation process, SE and T&E communities are expected to preserve the existing inextricable link with the RBT&E test approach and thus forecast SOI capability (vice demonstrate due to unavailability of integrated prototypes) through interpretation and analysis of available RBT&E-collected verification data. There exist two methods to do so. First is the standalone analyses and quantitative amalgamation of available RBT&E element verification data to interpret their collective effects (Bjorkman,

Gray, & March, 2005). Second is the aggregate and coherent simulation of respective Model-Based Systems Engineering (MBSE) low-level models (previously developed during the SOI Decomposition and Definition processes) utilizing the available RBT&E element verification data as model inputs (Eppinger, Whitney, Smith, & Gebala, 1994; Estefan, 2008; Ramos, Ferreira, & Barceló, 2013). To date, it is difficult to execute either of these two SOI capability-forecast methods during the SOI Implementation process because the available RBT&E element verification data present the following challenges:

1) Mismatch of element variables' units of measurement (or data metrology inconsistency): with each SOI element-under-test being developed under different requirements, their respective RBT&E verification data are collected and represented by unique variables and units of measurements specific to only their allocated requirements (e.g., *mph*, *tons*, *Watts*, etc.). This data-metrology inconsistency among the RBT&E element verification data prevents their subsequent mathematically- and statistically-conforming amalgamation, or aggregate simulation, to produce the corresponding SOI-capability forecast (Fiebrandt & Wilson, 2008).

2) Mismatch of element variables' measurement conditions (or data-domain inconsistency): with each element requirement specifying different verification conditions (e.g., at $x_1\ ^oC$ ambient condition, at x_2-*meters* altitude, after x_3-*hours* of operation, etc.), their respective RBT&E verification data are collected and presented as unique scalar-conditional variables. This data-domain inconsistency among the RBT&E element verification data violates the amalgamation property's joint embedding signature requirement (Hodges, 1997) and prevents their

subsequent linear amalgamation, or coherent simulation, to forecast the corresponding SOI capability as confirmed by Helle & Schamai (2016), Shaikh & Moore (1997), and Ogren (2000).

RBT&E's data-metrology and -domain inconsistency challenges result in the SE and T&E communities reverting to either: (a) subjectively (or judgmentally) interpret RBT&E element verification data and their collective effects (Bjorkman, et al., 2013), and deliver imprecise and/or incomplete SOI-capability forecasts that may hinder factual differentiations of risks from uncertainties (Parry, 1996; Corral, 1997); or (b) rely on synthetic element performances (e.g., respective element requirements' thresholds or historic manufacturer engineering data) to continue to deliver (what is now routine, obsolete, and pre Implementation process) SOI-design inspections vice SOI capability forecasts per Ogren (2000); or (c) just entirely overlook the *SOI-Capability Identification* T&E goal within the SOI Implementation process and deferring to subsequent integrated-level verification and validation (V&V) processes.

This study aims to resolve RBT&E's data-inconsistency challenges and SOI Implementation process' inability to objectively forecast SOI capability with its developed Scenario-Based Experimental-Design test approach (SBED).

1.2 Statement of the Problem

Despite the ever pressing need to identify SOI-level capability as early as possible within SOI development life cycle to further support assessment (and mitigation) of potential technical risks and uncertainties as pertain to the SOI design, the focus of the SOI

Implementation process continues to fixate exclusively on the capturing and collection of element performances and verification data specific only to the allocated element requirement variables and measurement conditions (i.e., employment of the traditional RBT&E test approach) as stated in INCOSE (2015) and DoD (2017). While the implementation of the RBT&E test approach introduces significant benefits toward the objective verification of element performance-under-test's requirement compliance (i.e., the accomplishment of the *Requirement-Based Performance Verification* T&E goal), its collected element verification data are represented by unique measurements and measurement conditions that bear minimal relevance to the overall integrated SOI capability (Helle & Schamai, 2016; Bjorkman &Gray, 2009). Thus, it is a challenge to quantitatively amalgamate, or coherently simulate, these mismatched RBT&E element verification data to produce an objective SOI-capability forecast.

To support the SOI Implementation process' ability to accomplish the *SOI-Capability Identification* T&E goal, there exists a need to resolve RBT&E's existing data-metrology and -domain inconsistency challenges. RBT&E's bottom-up reductionist verification approach needs to be supplemented (not replaced) with a top-down SOI-synoptic verification perspective. All individual (previously mismatched) element performance data need to be characterized as variables that are: (a) standardized w.r.t. a common unit-of-measurement that is relevant to the SOI mission; and (b) normalized w.r.t. a common domain that is comprehensively representative to the SOI operational environment. Thus, introducing to the SOI Implementation process with the ability to quantitatively amalgamate or coherently simulate these (now matching) element

performance variables to objectively forecast the SOI operational capability. And this is otherwise not available until subsequent Integration and Verification processes.

The author claims such a supplement to the RBT&E test approach exists in this developed SBED test approach.

1.3 Approach

This study researches and introduces the SBED test approach to supplement (not replace) the traditional RBT&E approach. SBED resolves the SOI Implementation process' existing RBT&E data-metrological and -domain inconsistency challenges by capture and characterize all individual element performances data as response models w.r.t. a common SOI mission-relevant unit-of-measurement (i.e., *SOI-element capability,* or the ability to satisfy allocated element requirement's specified threshold) and as influenced by a common SOI mission-representative domain (i.e., SOI's intended operational environment). SBED then either quantitatively amalgamate or coherently simulate the (now matching) element models to objectively produce the corresponding SOI-capability forecast. As illustrated in Figure 1-1 and further discussed in Subsection 3.2 (SBED Methodology Description), SBED is implemented through the following 5 steps:

Step 1. Depicts (or "scenarioizes") the SOI capability through a graphical representation of interacting and sequenced SOI's child elements (e.g., an entity relationship (ER) diagram, and other Functions-Based Systems Engineering (FBSE) or MBSE products, etc.).

Step 2. Introduces the parameter(s) that define SOI's intended operational environment as the common input domain to normalize all to-be-collected element verification data.

Step 3. Employs design of experiment (DoE) test methods to collect all element verification data and capture the effects (on the element functions-under-test) from changes within the SOI operational environment.

Step 4. Through Response Surface Methodology (RSM) statistical and analytical techniques, models all the collected element verification data as individual *element-capability* response models as influenced by the common SOI operational environment and standardized w.r.t. a quantitatively calculated *performance-versus-requirement* threshold response variable.

Step 5. Either: (a) Boolean-algebraically sums (i.e., super positions) all *element-capability* response models, w.r.t. the SOI-capability scenario developed from SBED's Step 1, to quantify their net response (or collective effects) and objectively produce the SOI-capability forecast model that is also influenced by the SOI operating environment; or (b) delivers all standardized and normalized *element-capability* response models to MBSE as their respective low-level models' inputs to aggregate coherent simulate a similar corresponding SOI forecast.

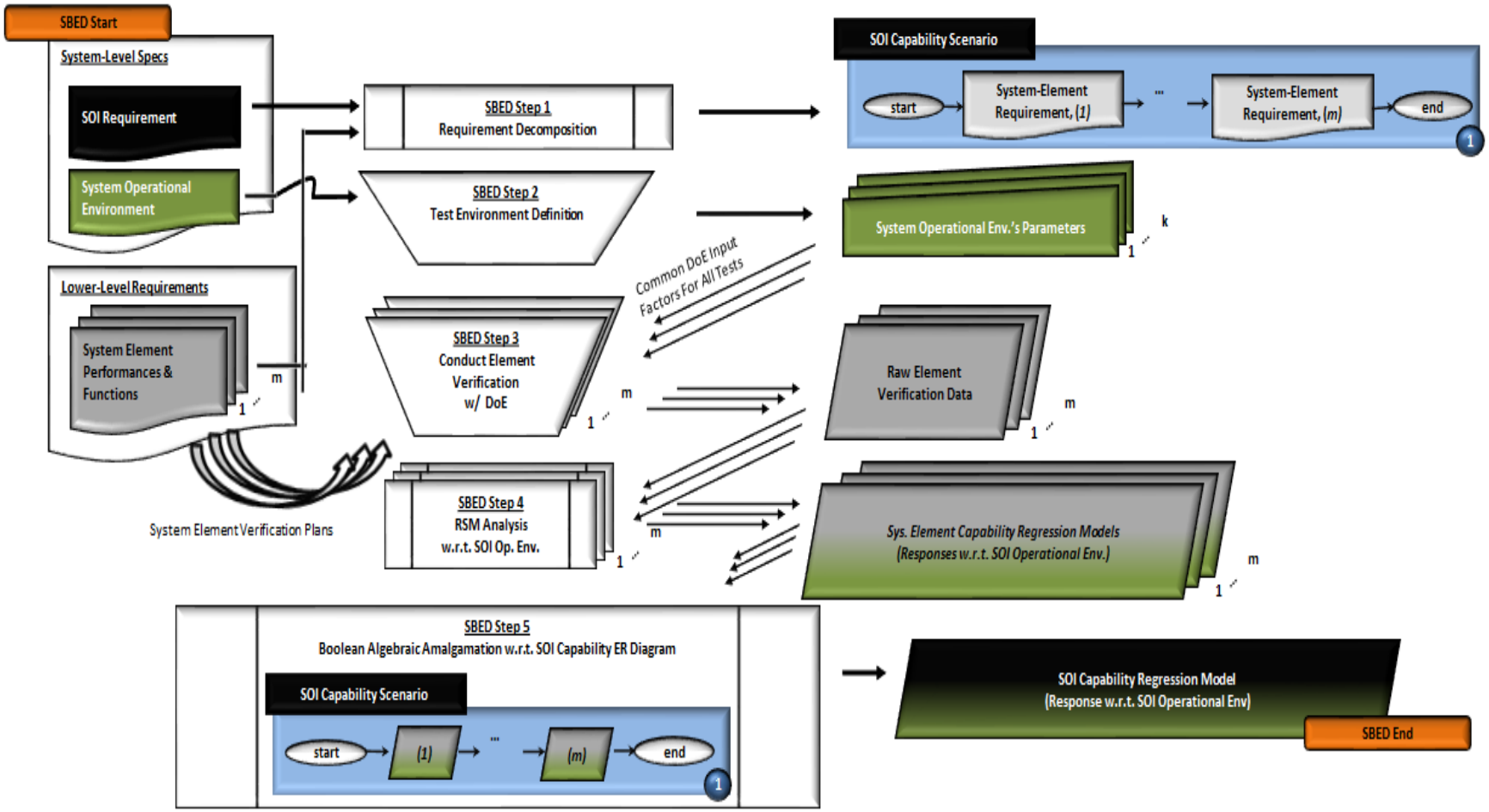


Figure 1 - 1, The SBED Test Approach's Model Diagram and Data Flow Chart

1.4 Organization of the Document

This dissertation is comprised of 5 chapters. Chapter 1 is an introductory providing background, purpose, scope, and goal on the subject matter while detailing the overall problem being considered. Parts of Chapter 1's content has been accepted in an upcoming IEEE Systems Engineering Journal article (Li, et al., 2018a). This article has completed peer review and accepted by the journal in May of 2018 but is not yet published at the time of this dissertation release. Chapter 2 presents an in-depth literature review of the traditional *Requirement-Based Performance Verification* and the new *SOI-Capability Identification* T&E goals, as well as the SOI Implementation process' challenges encountered towards the concurrently accomplishment both T&E goals through employment of the traditional RBT&E test approach. Chapter 2 also describes several existing RBT&E-supplemental T&E approaches and methods (including the early prototype T&E approach and MBSE's aggregated coherent simulations) aimed to identify and forecast integrated SOI-level capability, along with each of their specific benefits, limitations, and research gaps. All of Chapter 2's content is related to research and literature from both Li, et al. (2018a) and Li, Roberts, & Grenn (2018b). The latter is an upcoming INCOSE's Systems Engineering Journal article which is currently undergoing peer review at the time of this dissertation release. Chapter 3 presents a detailed description of this study's proposed SBED test approach. Chapter 4 presents a data-analysis example which further demonstrates SBED's ability to resolve RBT&E's data-metrological and –domain inconsistency challenges and to objectively forecast SOI capability (i.e., normalize, standardize, and amalgamate element verification data) by applying its DoE and RSM analytical techniques onto an existing United States Air Force (USAF) Flight Test Center

(AFFTC) T-38C flight test data set (Gunn-Golkin et al., 2010). This data-analysis example, as initially introduced in (Li, et al., 2018a), further exhibits SBED-introduced SE and T&E benefits by comparing SBED's quantitatively-produced capability forecast against AFFTC's original qualitative SOI assessment conclusions (as subjectively drawn from standalone analyses of individual RBT&E-collected element verification data). Chapter 5 provides conclusions and recommendations for future research.

1.5 Significance

SBED introduces several SE and T&E benefits to the SOI Implementation process. The first introduced benefit of SBED is the early delivery of an objective and quantitative SOI operational-capability forecast. This is previously unobtainable during the SOI Implementation process, to accomplish the *SOI-Capability Identification* T&E goal (Li, et al., 2018a).

The second introduced benefit applies directly within the SOI Implementation process's ability to accomplish the *Requirement-Based Performance Verification* T&E goal. It is SBED's deliveries of more detailed and more SOI-mission representative *"element-capability"* models from SBED's Step 4 as discussed in Subsection 3.2.4 (Calculate Standardized *Element-Capability* Models). SBED's *element-capability* models comprehensively characterize every respective element performance w.r.t. the entire SOI operational environment; thus, expanding the scope of Implementation process' primary *Requirement-Based Performance Verification* T&E goal (which has been previously not expandable) to beyond just the element requirement-specified verification conditions (Bjokman & Zeis, 2007; Eldh, Brandt, Street, Hansson & Punnekkat, 2010; Hsueh, 1995).

Additionally, the provision of SBED's more detailed and more SOI-mission representative *"element-capability"* models further improves the Implementation process' ability to objectively validate MBSE's corresponding design models and support MBSE's aggregated coherent simulation of corresponding low-level models to forecast SOI capability. Together, the SBED-produced SOI capability forecast and its corresponding *element-capability* models objectively inform SOI decision makers and designers of potential risks and uncertainties associated to the SOI design as well as the identification of culpable child element(s) (Li, et al., 2018a).

Overall and as illustrated in Figure 1-2, the supplementation of SBED's top-down SOI-synoptic verification approach establishes the Implementation process' SE-information feedforward to future Integration and Verification processes along the SE-Vee as defined in Forsberg and Mooz (1991) by providing objective means to both understand the SOI emergent behaviors w.r.t. the SOI hierarchical structure (Anandalingam & Friesz, 1992) as well as to identify subsequent integrated test subjects (Zhu, Pritchard, & Silverberg, 2016). Concurrently, the supplementation of SBED also establish the Implementation process's SE-information feedback to previous adjacent SOI design Decomposition and Definition processes along the SE-Vee by providing objective means to understand, inspect, and validate the existing SOI design and SOI-design products (Gotel & Finkelstein, 1994). For example, corresponding MBSE developed lower-level models.

All SBED-introduced benefits are discussed in Subsection 3.3 (Additional Benefits of SBED) of this dissertation and further demonstrated through Chapter 4's data-analysis example.

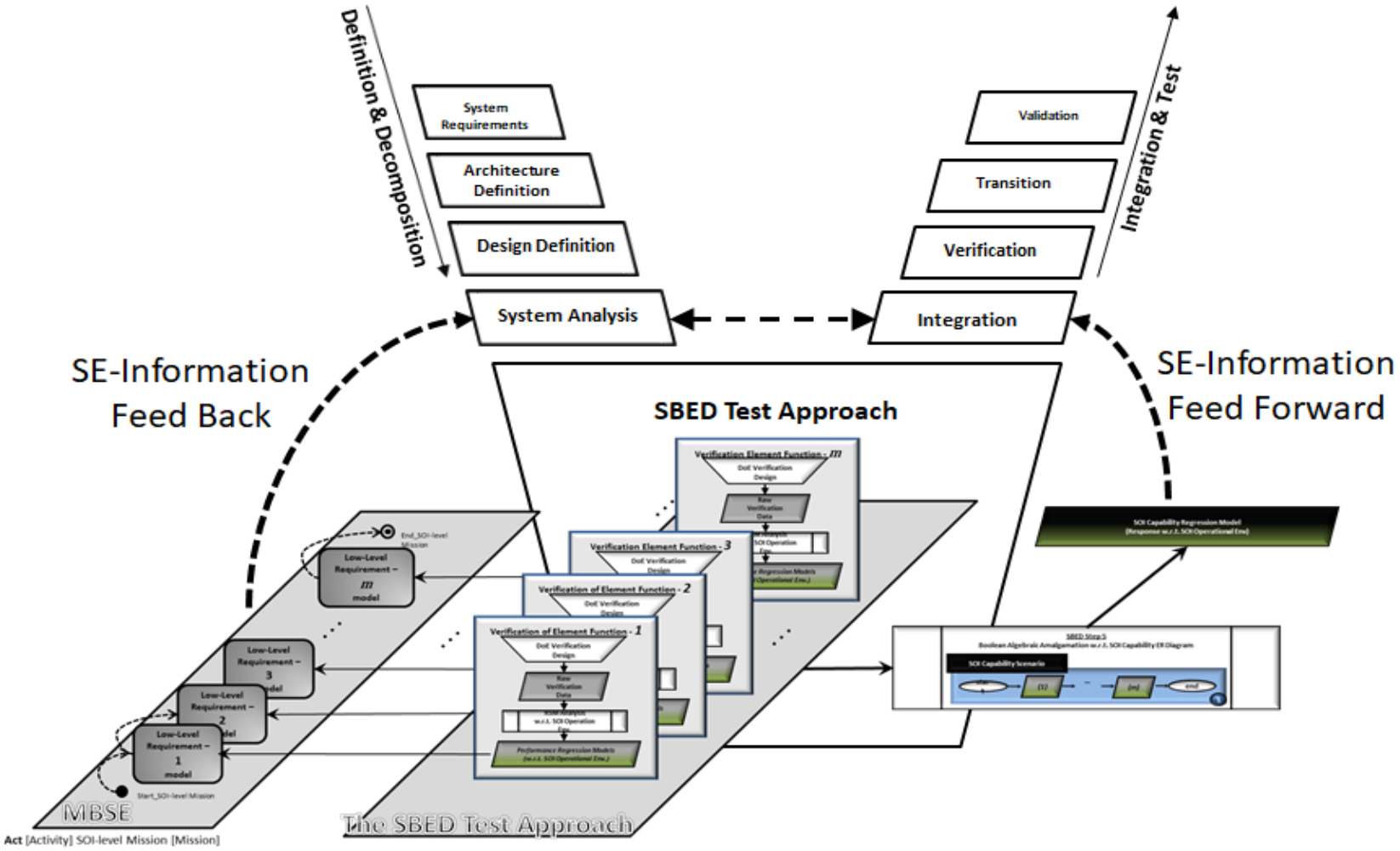


Figure 1 - 2, SBED Information Feedforward- and Feedback Flow within the SE Vee

1.6 Scope and Limitations

Multiple existing SE-based T&E frameworks and approaches have been introduced and employed throughout the SOI Implementation and Integration-and-Verification processes intended to introduce a top-down SOI-synoptic verification approach. As further discussed in Section 2.3 (Existing RBT&E Supplements to Forecast SOI Capability) of this dissertation, most of these existing frameworks and approaches are aimed to address a variety of current SE- and T&E-related programmatic-management challenges and difficulties specific to software-centric system's development. In comparison, this study's proposed SBED test approach presents three unique distinctions in its scope:

1. SBED's primary goal is to introduce the ability to objectively forecast SOI-operational capability. This study places minimal emphasis towards improving the resource-management and -allocation aspects of the SOI development life cycle;
2. SBED is primarily intended for the SOI Implementation process and its element verification activities. While SBED can be re-employed during subsequent SOI integration and test activities (i.e., its SOI-hierarchical modularity benefit as discussed in Subsection 3.4.4.1), the author defers the research of SBED's utilities and benefits toward the later Integration and Verification processes as future works; and
3. SBED is tailored towards hardware-centric SOIs. This is because: (a) software-centric SE and T&E frameworks and approaches are already a well-research topic, and (b) existing software-centric T&E frameworks and approaches differ from their traditional hardware-centric counterparts as a heavy reliance is placed

on the availability of higher verification data sample sizes. This is a luxury that is typically unavailable to hardware-centric SOI V&V activities, especially during the initial Implementation process.

The methodology of this study's SBED test approach places a heavy reliance on incorporating DoE test methods. This study emphasizes that its intent is not to teach specific DoE designs or DoE-design requirements during SOI Implementation process' element verification activities. While SBED is dependent upon the employment of DoE designs to capture individual element performances w.r.t. the SOI operational environment, the actual selections of appropriate DoE designs to be specifically employed during that SOI's element verification test remain to be unique case-by-case decision that are based on each element verification's test-resource availabilities and its intrinsic *Requirement-Based Performance Verification* objectives. This emphasis is further iterated in Subsection 3.2.3 (Design and Execute Element Verifications with DoE Method) of this dissertation.

Furthermore, this dissertation includes a data-analysis example to demonstrate the utilities and benefits of SBED that is presented with limited scope. The example utilizes a real-life USAF flight-test data set to partially demonstrate and validate SBED's data-analysis methodology. However, it does not offer a complete (start-to-end) and realistic demonstration of SBED methodology because a complete real-life data set that can fully exhibit this study's problem statement is not available for this study. Therefore, a comprehensive example utilizing a complete, real-life, and SOI hierarchical data set (implying that includes quantitative performance data at both the system-element level and the corresponding integrated SOI level) is listed as future work in Chapter 5.

Finally, because the propose SBED test approach places a minimal emphasis on the schedule and resource aspects of T&E, as well as the overall SOI development life cycle, the delivery of SBED-produced SOI- and element-level capability forecasts cannot solely dictate the outcome of SOI designers and decision maker's responses to immediately implement any respective corrective actions. The goal SBED test approach is to produce comprehensive and objective element- and SOI-level capability forecasts to support the assessment and mitigation of any potential technical risks associated to the SOI design. However, with the SOI Implementation process taking place after what are typically long and arduous SOI Decomposition and Definition processes, the overall significance and influence of any SBED-identified and -reported technical risks and uncertainties may be attenuated when equal amounts of considerations towards other SOI development's programmatic constraints are in place (i.e., schedule, cost, and resources). Therefore, there remains a need to development and instill a SOI-centric T&E approach during the earlier SOI Decomposition and Definition process to minimize any corrective actions' potential programmatic impact. This need is listed as future work in Chapter 5.

2 Literature Review

2.1 The Traditional Roles and Expectations of T&E

2.1.1 The Requirement-Based Performance Verification Goal

During the SOI Implementation process, the paramount T&E goal is to confirm that the elements are "built right" through inspections, verifications, and documentation of captured performances against their respective allocated requirements. This is the *Requirement-Based Performance Verification* T&E goal. The successful accomplishment of the *Requirement-Based Performance Verification* T&E goal requires implementing individual well-defined and well-planned test events and artifacts that are focused exclusively to the implemented or integrated functions under test (INCOSE, 2015; DoD, 2017; Ministry of Defense [MODAF], 2005) and relying on the corresponding collected performance verification data to objectively differentiate the function-under-test's intended behavior (i.e., requirement thresholds) against actual demonstrated behavior (Helle & Schamai, 2016).

For standalone verifications of implemented SOI-element functions, its corresponding captured verification data must be directly pertinent only to the allocated element requirement's specific performance metrics (Weiss & Schmidt, 2010). For verifications of SOI-level (or integrated-level) functions, it becomes more complicated. Given the sheer number of potential decomposed element requirements within an SOI design (e.g., a typical planned V&V portfolio for a major DoD weapon system may include dozens or even hundreds of individual verifications which also includes hundreds or thousands of individual test cases and scenarios according to Bjorkman, Sarkani, & Mazzuchi (2013), its corresponding captured verification results must be comprehensively

pertinent to all combinations of operational input parameters as collectively specified within all applicable element- and integrated-level requirements per Shaikh & Moore (1997).

As T&E typically assuming the role of being the byproduct of earlier SOI Development Process schedule slides and budget cuts, there exist expectations of accomplishing the primary *Requirement-Based Performance Verification* goal with less resources and time to absorb the already-incurred negative programmatic impacts (Hess & Valerdi, 2010; Harman, Cortez, & Hill, 2016). This is especially the case during the SOI Implementation process, of which is the initial T&E activities encounter after the SOI Decomposition and Definition processes. As a result, it is often considered prohibitively resource intensive to adequately verify integrated SOI-level functions, especially w.r.t. the "infinite" number of nonhomogeneous parameters as pertinent to all the applicable element requirements' specified configuration and operating conditions (Souag, Mazo, Salinesi, & Comyn-Wattiau, 2016; Fiebrandt & Wilson, 2008; Bjarnason, et al., 2014). Thus, to maintain the assurance that the captured verification results can still objectively inform SOI designers and stakeholders of potential uncertainties and risks associated to the selected design approach (DoD, 2017; Bjorkman, et al., 2005), the Implementation process' most pragmatic solution is to remove the cumbersome and expensive verifications of integrated SOI-level functions, and reduce the focus and scope of its performance verification down to the most fundamental and lowest SOI-hierarchical level: standalone verification of individual element functions only against their allocated element requirements (e.g., only demonstrate y_1-mph speed under $x_1{}^{o}$C verification condition, only demonstrate y_2-tons of payload at x_2-m altitude, and only demonstrate sustainment of y_3-watts of output power

after x_3-hours of operation, etc.) as remarked in Corral (1997) and Ncube (2011). This bottom-up reductionist verification approach (of fixating exclusively on the capturing and collection of element performances and verification data specific only to the allocated element requirement variables and measurement conditions), by definition, is the traditional RBT&E test approach. And it is inextricably linked to the SOI Implementation process.

The RBT&E test approach is the most fundamental, most practical, and most objective SOI-element verification methodology to employ during the SOI Implementation process (Hesari, et al., 2013). This is because its methodology focuses exclusively on generating individual and standalone test artifacts that are specific only to the allocation requirement's performance metrics and verification conditions (INCOSE, 2013; Bjorkman & Gray, 2009) and its goal exclusively includes only the early and reliable up-front identification of element-level design defects (INCOSE, 2015; Helle & Schamai, 2016). RBT&E completely removes the top-down SOI-synoptic perspective of assessing element performances as they pertain to the SOI-level capabilities and defers all integrated SOI-level requirement verifications to subsequent Integration and Verification processes (Eigbe, et al., 2010). As further discussed next, RBT&E bottom-up reductionist verification approach's concomitance to the SOI Implementation process's nature (and its external programmatic demands) establishes an inextricable link between the two.

2.1.2 Benefits of The Traditional RBT&E Approach

To the best knowledge of this study, the implementation of RBT&E and its associated challenges toward objective performance-defect identification (and ultimately

accomplishment of the *Requirement-Based Performance Verification* T&E goal) are rarely reported in industry-related journals (Bjorkman, et al., 2013; Romas, et al., 2013; Zhu, et al., 2016). This is because RBT&E's methodology placates to the fundamental nature of SOI Implementation process, as well as introduces several immediate SE and T&E benefits. These benefits are further discussed I the following subsections.

2.1.2.1 Adaptability Towards SOI Design's Hierarchical Complexity

The first and foremost benefit of the RBT&E test approach, towards accomplishing the traditional *Requirement-Based Performance Verification* T&E goal, is its ability to objectively differentiate between the individual element function-under test's intended behaviors (i.e., allocated requirement thresholds) and its actual behavior, or to gain confidence towards the lack of thereof (Helle & Schamai, 2016; Parry, 1996; Bjorkman & Gray, 2009). RBT&E's methodology, of generating isolated test artifacts specific only to the individual element functions-under-test, quarantines its collected element verification data from the influences of potential interactions from other element functions (Bjorkman, et al., 2013). This allows the subsequent interpretation an characterization of captured element verification data to: (a) Provide full insights of the element design and implementation as conveyed through the boundaries of the allocated requirement descriptions (e.g., what information, who created it and when, where it came from, how it is presented, and why) as described in Winkler & von Pilgrim (2009) and Ramesh & Jarke, (2001); and (b) Factually differentiate between issues, risks, and uncertainties associated to the respective element's design and implementation that may ultimately affect stakeholders' decision outcomes (Shaikh & Moore, 1997; Parry, 1996; Corral, 1997).

RBT&E also reduces the complexity of generating element-verification test artifacts by quarantining its element-level test-design and -planning efforts from any integrated SOI-level considerations (Bjorkman, et al., 2013). Typically during the SOI Implementation process, corresponding SOI-level user requirements still remain to be assumptions and implicit interpretations of solicited colloquial, ambiguous, and subjective SOI user statements (Liu, Zhang, & Chen, 2012), and their respective and applicable element designs (that spread over different engineering disciplines) are often individually created with little to no consideration of the effect to the overall SOI-level user requirement (Corral, 1997; Engle, Bogomolni, Shachar, & Grinman, 2004). Therefore, as one single element function combines with tens and hundreds of others (i.e., as into one black box that is only defined by required inputs and outputs), the complexity in generating the appropriate test artifacts would significantly increases while corresponding testability decreases (Dahmann, Lane, Rebovich, & Lowry, 2010; Apostolov & Vandiver, 2014). In return, RBT&E reduces the complexity of Implementation process' verification design and planning efforts by ignoring the hierarchical relationship between SOI-level requirements and its corresponding element requirements and ultimately eliminating the need to comprehensively demonstrate element performances w.r.t. to the "infinite number" of combinations of operational input parameters as specified to the all allocated SOI and element requirements (Bjorkman, et al., 2013).

Furthermore, RBT&E offers a level of robustness against the constant (and often vague and ambiguous) requirement additions, deletions, and modifications that are ubiquitous to the dynamic SOI development life cycle (Corral, 1997; General Accounting Office [GAO], 2004; DoD, 2010). These 'requirement volatilities', as characterized by

Peña & Valerdi (2015), are often individually introduced at the element-requirement level and of which the ultimate effects to remaining element designs remain unknown and often unconsidered (Dahmann, et al., 2010; Loren, 2007). Therefore, RBT&E's focus of generating only element requirement-specific test artifacts advantageously allows SOI Implementation process' verifications to ignore both the additional complexity introduced by this dynamic nature of SOI design and the potential challenges encountered toward the partitioning organic element performance defects from the influences and effects of other element functions' potential interactions (Di Nitto, Ghezzi, Metzger, Papazogou, & Pohl, 2008; Ncube, 2011).

2.1.2.2 Suitability to Independent & Asynchronous Element-Development Strategy

The RBT&E methodology placates to the typical schedule- and geographically-disparate nature of independent and asynchronous SOI-element development strategy and their respective verification requirements (Li, et al., 2018a).

SOI element designs are typically independently executed as autonomous entities starting from different levels of initial readiness and maturity, and asynchronously delivered for respective performance verifications with minimal constraint from SOI-level verification-planning considerations (INCOSE, 2015; Dahmann, 2015). Additionally, individual element verification test cases and scenarios often mandate unique capabilities from geographically disparate test facilities and supply chains (e.g., extreme environmental condition test, electromagnetic interference and compatibility test, live-fire lethality test, etc.) as emphasized in DoD (2003). Therefore, it is difficult (if not impossible) to

synchronize and collocate all individual element test artifacts during the SOI Implementation process (Dahmann, et al., 2010). RBT&E's scope, of executing and verifying only element requirement-specific test events, placates to this schedule- and geographically-disparate nature of SOI-element verification (Weiss & Schmidt, 2010).

2.1.2.3 Additional SE and T&E Benefits

RBT&E methodology also introduces several intrinsic SE and T&E benefits to the Implementation process' accomplishment of the *Requirement-Based Performance Verification* T&E goal.

One is the establishment of clear and bidirectional traceability between individual element verification results (as well as their identified element performance defects) and their allocated element requirements (INCOSE, 2015). With RBT&E's methodology of generating only element requirement-specific test artifacts, captured element performance and identified defects can be directly correlated to the allocated element requirement under test. This provides traceability on how the identified performance defects, and its probability of occurrence, may potentially impact the integrated SOI design (albeit subjectively) based on the priority of the respective element requirement (e.g., key performance parameters, critical technical parameters, and measures of performance) as pointed out in Ncube (2011).

Another intrinsic SE benefit of RBT&E is its ability to validate the element requirements while concurrently verifying element performance w.r.t. that requirement (INCOSE, 2015). With RBT&E scope of generating isolated test artifacts specific only to the individual element functions-under-test, the quality of respective test cases and test scenarios can be utilized as qualitative reflection to gauge the allocated element

requirements' organic attributes (i.e., is the requirement's specified measurements and verification conditions testable, measurable, quantifiable, etc.) as discussed in Winkler & von Pilgrim (2009) and Graham (2002). Vague, ambiguous, or unclear requirements can be objectively identified and corrected during RBT&E's test artifact-design and -planning efforts during the SOI Design and Decomposition processes (Graham, 2002; Zhang, Li, & Liu, 2014). Prior to the actual start of an Implementation process' element verification activities.

2.1.3 Implementation Process' Inextricability Towards RBT&E

With all of its intrinsic SE and T&E benefits towards accomplishing the *Requirement-Based Performance Verification* T&E goal, the RBT&E test approach provides a feasible, reliable, and robust method to objectively verify that SOI elements-under-test perform their intended functions and fulfill their allocated performance requirement thresholds and characteristics. Therefore, it remains to be inextricably linked (and justifiable so) to the SOI Implementation processes per Li, et al. (2108a).

However, there now exists a need to expand the focus and scope of the SOI Implementation process' element verifications to concurrently assess element performances as pertinent to the integrated SOI-level mission and across a wide range of possible operational environments as stated in Bjorkman & Gray (2009) and Eigbe, et al., (2010). The SOI Implementation process must now concurrently address these challenges associated to SOI-level assessments that RBT&E's methodology has previously made possible to ignore (e.g., complexity of SOI hierarchy, independent and asynchronous element development).

2.2 New Roles and Expectations of T&E

2.2.1 The SOI-Capability Identification Goal

While this *Requirement-Based Performance Verification* T&E goal remains paramount during element verification activities, there now exists another equally important T&E goal to identify SOI operational capability (or the next integrated level's capability) in support of the identification and mitigation of risks associated to the implemented design (Eigbe, et al., 2010; Helle & Schamai, 2016; Li, et al., 2018a). This is the new *SOI Capability-Identification* T&E goal.

The new *SOI Capability-Identification* T&E goal coincides with the early SOI Decomposition and Definition processes' top-down SOI-synoptic practices which fixates its primary design focus on the SOI-level user requirements as details of individual aspects of an SOI design are developed (INCOSE, 2015; Pyster, et al., 2012;, Unterkalmsteiner, Feldt, & Gorschek, 2014; Anandalingam & Friesz, 1992; van Lamsweerde, Darimont, & Leiter, 1998). To maintain the alignment, the V&V activities throughout subsequent SOI Integration and Verification processes are also expected to incorporate the same top-down SOI-synoptic verification approach (Ncube, 2011) and collect corresponding performance test results to: (a) Determine if the implemented functions-under-test solve the SOI-level user requirement through associations and correlations of identified potential uncertainties, risks, and potential technology gaps to the overall SOI-design approach (DoD, 2003; Bjorkman, et al., 2013); and (b) Identify the next integrated-level SOI's emergent behavior to optimize identification of subsequent Integration and Validation process' test subjects (INCOSE, 2015; DoD, 2017; Zhu et al., 2016).

As a large portion of total SOI life cycle's cost is committed by decisions made within its development stage (Loren (2007) states eighty percent), the implementation of a top-down SOI-synoptic verification approach to accomplish this *SOI Capability-Identification* T&E goal is expected to take place as early as possible to achieve performance and minimize cost (Roberts, Mazzuchi, & Sarkani, 2016) and the initial opportunity occurs during the Implementation process and its element verification activities.

2.2.2 Inadequacies of the RBT&E Approach

During the SOI Implementation process, the *SOI-Capability Identification* T&E goal is expected to be accomplished without breaking the existing inextricability towards RBT&E and RBT&E's benefits despite that the fact that they are applicable only toward accomplishing the primary *Requirement-Based Performance Verification* T&E goal. Therefore, the corresponding SOI capability must be forecasted via a top-down SOI-synoptic approach either through the quantitative amalgamation of available RBT&E-collected element verification data, or through the coherent simulation of respective low-level MBSE models using the same RBT&E data as respective low-level model inputs (Bjorkman & Gray, 2009). To date, it is a challenge to implement either of these two SOI-capability forecasting methods during the SOI Implementation process. This is because traditional RBT&E-collected element verification data (e.g., only demonstrate y_1-mph speed under x_1°C verification condition, only demonstrate y_2-tons of payload at x_2-m altitude, and only demonstrate sustained y_3-watts of output power after x_3-hours of operation, etc.) intentionally exclude this needed top-down SOI-synoptic verification

perspective. This presents two RBT&E's data-inconsistency challenges towards the objective forecasting of the SOI capability (Li, et al., 2018a).

The first challenge pertains to arithmetical nonuniformity. With SOI elements being developed under different requirements, their respective RBT&E-collected verification data are represented by different variables and different units of measurement specific to only their allocated requirements (e.g., *mph*, *tons*, *Watts*, etc.). These element verification variables are inconsistent to each other and also most likely to the SOI-level requirement's performance measurements. This data-metrology inconsistency (or absence of uniformity in units of measurement) of RBT&E element verification data prevents any integrated SOI-level capability forecast to be mathematically- and statistically-amalgamated (Fiebrandt & Wilson, 2008).

The second challenge pertains to nonlinearity in variable domains. With each element requirement specifying different verification conditions (e.g., at $x_1{}^oC$ ambient condition, at x_2-*m* altitude, after x_3-*hours* of operation, etc.), their respective RBT&E-collected verification data are presented as mismatched and scalar conditional variables. Furthermore, each element requirement's verification condition typically does not fully represent the actual operational environment of the respective element design nor the integrated SOI's mission (Hsueh, 1995). This data-domain inconsistency challenge (or the absence of a common and SOI-mission representative measurement conditions) violates the amalgamation property's joint embedding signature requirement (Hodges, 1997) and prevents their linear and quantitative amalgamation to produce an SOI mission-representative capability forecast (Helle & Schamai, 2016; Shaikh & Moore, 1997).

Together, RBT&E's element data-inconsistency challenges result in the SOI Implementation process reverting to either subjectively (or judgmentally) interpret element verification data (and their collective effects) and outputs imprecise and/or incomplete SOI capability forecasts that may hinder factual differentiation of risks from uncertainties (Parry, 1996; Corral, 1997), or just entirely overlook the *SOI-Capability Identification* T&E goal throughout the process's element-verification activities.

2.2.3 Inadequacies to Support MBSE's SOI-Capability Simulation

Throughout the SOI Decomposition and Definition processes, MBSE has become a standard SE practice to inspect the SOI design quality requirement by requirement, as further discussed by works by Karban, et al. (2008), Ramos, et al. (2013), and Kalawsky, et al. (2013), Crisp (2007), Topper & Horner (2013), and Holt, et al. (2015). It has emerged as a formalized and cost-effective SE tool to identify and resolve any potential complexity and uncertainty issues associated to the SOI design based on the experimental simulations of SOI's multifaceted low-level requirements (Corral, 1997; Chen, et al., 2012; Ramos, et al., 2013; Kernstine, 2013; Toper & Horner, 2013). Once the SOI development life cycle transitions (from design) to its Implementation process, MBSE's role as an SE tool is also expected to transition from SOI-design inspection to SOI-design exploration. It is now expected to identify the corresponding integrated SOI-level's emergent behavior, w.r.t. the domain of its intended operational environment (i.e., the *SOI-Capability Identification* goal), through the aggregated simulation of available RBT&E element verification results as its respective low-level model simulation inputs (Winkler & von Pilgrim, 2009;

Eppinger, et al., 1994; Estefan, 2008; Ramos, et al., 2013; Morandini, Penserini, Perini, & Marchetto, 2017).

However, during the SOI Implementation process, MBSE's ability to forecast SOI capability encounters the same RBT&E's data-domain inconsistency challenges. MBSE cannot objectively identify an SOI capability, one that is representative of the SOI-mission and comprehensive to the SOI-operational environment, unless the available RBT&E element verification data utilized as simulation inputs also bear the same characteristics (Winkler & von Pilgrim, 2009; Shaikh & Moore, 1997; Fiebrandt & Wilson, 2008). With the knowledge that such extension of the RBT&E test approach and data are not feasible (Bjorkman & Zeis, 2007), MBSE reverts to relying on historic and synthetic element performances (e.g., respective low-level requirements' thresholds or manufacturer's engineering data) to continue to produce (what is now routine, obsolete, and pre-Implementation process) SOI-design inspections during the SOI Implementation process. Its identification of SOI capability are deferred until subsequent integrated V&V results become available (Fiebrandt & Wilson, 2008; Boumen, et al., 2009; Ogren, 2000).

2.2.4 Supplementation of the RBT&E Approach

To demonstrate RBT&E's existing data-metrological and –domain inconsistency challenges, Table 2-1 lists four applicable element requirements of a hypothetical SOI architecture and their respective descriptions. Each element requirement in Table 2-1 is intentionally introduced with unique requirement-specified performance-measurement variables and verification conditions (as illustrated by different Greek alphabet letters as unit variables and distinct variable designations, x_i^*s). Concurrently, Table 2-2 lists the

corresponding element verification data of Table 2-1's four requirements, as reported in formats commonly associated to the traditional RBT&E test approach. It is self-evident through examination of Table 2-2's RBT&E-collected element verification data, that due to each requirement's unique measurements, thresholds, and verification conditions, collectively they are inadequate to support any subsequent mathematically- or statistically-conforming amalgamations or coherent simulations to forecast the corresponding SOI-capability. Even if all of four requirements' verification test results are uniformly reported in one of the formats within Table 2-2 (i.e., any one of the three right-sided columns), as long as there exists one element requirement that specifies a unique unit-of-measurement or a verification condition, RBT&E's existing data-metrological and –domain inconsistencies present challenges towards the production of an objective SOI-capability forecast.

Table 2 - 1, Examples of Hypothetical Allocated Element Requirements

Requirement #	Threshold Values/Ranges	Threshold Units of Measurement [A]	Performance Specification Evaluation Criteria [A]	Verification Test Conditional Parameters [B, C]
1- *(f_1)*	$T_{i=1} \stackrel{def}{\cong} \{7.5 \text{ to } 15\}$	***A***	$7\,\boldsymbol{A} \leq \boldsymbol{f_1} \leq 15\,\boldsymbol{A}$	x_1*$\boldsymbol{B}$ and x_2*$\boldsymbol{\Gamma}$
*Corresponding allocated requirement: Low-level structure $\boldsymbol{f_1}$ must demonstrate performance between 7 and 14 (in units of **A**), under the specific condition as described by x_1* (in units of **B**) and x_2* (in units of **Γ**).*				
2- *(f_2)*	$T_{i=2} \stackrel{def}{\cong} 80$	***Δ***	$\boldsymbol{f_2} \leq 80\,\boldsymbol{\Delta}$	x_1*$\boldsymbol{B}$ and x_3*$\boldsymbol{E}$
*Corresponding allocated requirement: Low-level structure $\boldsymbol{f_2}$ must demonstrate performance of less or equal to 80 (in units of **Δ**), under the specific condition as described by x_1* (in units of **B**) and x_3* (in units of **E**).*				
3- *(f_3)*	$T_{i=3} \stackrel{def}{\cong} 75$	***E***	$\boldsymbol{f_3} < 75\,\boldsymbol{E}$	$x_{4(min)}\boldsymbol{Z} \leq \mathbf{x_4} \leq x_{4(max)}\,\boldsymbol{Z}$
*Corresponding allocated requirement: Low-level structure $\boldsymbol{f_3}$ must demonstrate performance of less than 75 (in units of **E**), under the specific conditions between $x_{4(min)}$ and $x_{4(max)}$ (in units of **Z**).*				
4- *(f_4)*	$T_{i=4} \stackrel{def}{\cong} 15$	***H***	$\boldsymbol{f_4} \geq 15\,\boldsymbol{H}$	x_5*$\boldsymbol{\Theta}$ and x_6*$\boldsymbol{I}$
*Corresponding allocated requirement: Low-level structure $\boldsymbol{f_4}$ must demonstrate performance of greater than 15 (in units of **H**), under the specific condition as described by x_5* (in units of **Θ**) and x_6* (in units of **I**).*				

Note A: Unit of measurement, for requirement thresholds and verification conditions, are intentionally presented in random Greek alphabets to illustrate the common inconsistencies among the system elements.

Note B: $\mathbf{x_{\#}}$ represents arbitrary and collectively non-uniformed verification conditions (i.e., nonmatching parameters).

Note C: $\mathbf{x_{\#}}$***** notation indicates a fixed constant value

Table 2 - 2, Reporting Formats of the Traditional RBT&E Verification Data

Requirement #	Test Conditions	Binomial Evaluation Results	Statistical Analysis Results [B]	Stochastic Results
			($\mu \pm 2 \cdot \sigma$)	# *of failures/total sample*
1- ***(f_1)***	x_1*$\boldsymbol{B}$, x_2*$\boldsymbol{\Gamma}$, x_3*$\boldsymbol{E}$ (combined)	$\bar{y}_{f_1}$ failures (N_{f_1} total)	($\mu_{f_1} \boldsymbol{\Delta} \pm \sigma_{f_1}{}^2$) | x_1*$\boldsymbol{B}$ and x_2*$\boldsymbol{\Gamma}$	$P(7.5\ \boldsymbol{A} \leq f_1 \leq 15\ \boldsymbol{A})$ | x_1*$\boldsymbol{B}$ *and* x_2*$\boldsymbol{\Gamma}$ = ($\bar{y}_{f_1}$ failures) / (N_{f_1} total)
2- ***(f_2)***		$\bar{y}_{f_2}$ failures (N_{f_2} total)	($\mu_{f_2} \boldsymbol{\Delta} \pm \sigma_{f_2}{}^2$) | x_1*$\boldsymbol{B}$ and x_3*$\boldsymbol{E}$	$P(\boldsymbol{f_2} \leq \boldsymbol{80\ \Delta})$ | x_1*$\boldsymbol{B}$ *and* x_3*$\boldsymbol{E}$ = ($\bar{y}_{f_2}$ failures) / (N_{f_2} total)
3- ***(f_3)***	$x_{4(min)}\boldsymbol{Z} \leq \boldsymbol{x_4} \leq x_{4(max)}\ \boldsymbol{Z}$	$\bar{y}_{f_3}$ failures (N_{f_3} total)	($\mu_{f_3} \boldsymbol{E} \pm \sigma_{f_3}{}^2$) | $\{x_{4(min)}\boldsymbol{Z} \leq \boldsymbol{x_4} \leq x_{4(max)}\ \boldsymbol{Z}\}$	$P(\boldsymbol{f_3} \leq \boldsymbol{75\ E})$ | $\{ x_{4(min)}\boldsymbol{Z} \leq \boldsymbol{x_4} \leq x_{4(max)}\ \boldsymbol{Z})$ = ($\bar{y}_{f_3}$ failures) / (N_{f_3} total)
4- ***(f_4)***	x_5*$\boldsymbol{\Theta}$, x_6*$\boldsymbol{I}$	0 failures (N_{f_4} total)	($\mu_{f_4} \boldsymbol{H} \pm \sigma_{f_4}{}^2$) | x_5*$\boldsymbol{\Theta}$ and x_6*$\boldsymbol{I}$	$P(\boldsymbol{f_4} \geq \boldsymbol{15\ H})$ | x_5*$\boldsymbol{\Theta}$ *and* x_6*$\boldsymbol{I}$ = *0*

Note A: All derivations are based on Table 2-1's arbitrary allocated requirement thresholds.

Note B: Prediction interval of ($2 \cdot \sigma$) is arbitrarily selected for this particular illustrative example.

Because of RBT&E's inability to collect matching element verification data that are needed to amalgamate, or to coherently simulate, an objective SOI-capability forecast, the Implementation process becomes inadequate to optimally identify subsequent process' test subjects, and the overall SOI Integration and Verification processes is typically one of the weakest and unorganized points along current SOI development life cycle (Bjorkman & Gray, 2009; Helle & Schamai, 2016). Despite SOI development life cycles devoting nearly 50% of its total effort towards element (or constituent system) performance verifications (Tsai, et al., 2010), only near half of the studied SOI development projects from Encontre (2003)'s and Helmerich, et al. (2005)'s studies are on schedule and of which respective SOI designs meeting performance and functionality expectations. As neither a bottom-up reductionist nor a top-down SOI-synoptic verification approach alone can provide a comprehensive understanding of the full hierarchy of the SOI design needed to concurrently accomplish both the *Requirement-Based Performance Verification* and *SOI-Capability Identification* T&E goals (Kline, 1995), the SOI Implementation process need to adopt a supplemental (not replacement) top-down SOI-synoptic verification approach to enable both: (a) the characterization of the element function-under-test w.r.t. SOI-mission relevant measurements across the full spectrum of operational environment (Corral, 1997; Ncube, 2011; Bjorkman & Gray, 2009); and (b) the concurrent preservations of RBT&E's existing bottom-up reductionist verification approach' rigor towards the *Requirement-Based Performance Verification* T&E goal and its benefits.

To date, it is difficult for the T&E community to present such a complete supplemental solution to the RBT&E approach during SOI Implementation process. Beyond having the T&E community to (first and foremost) actively participate during SOI

Decomposition and Definition processes' verification designing and planning activities (e.g., participating in requirement and design reviews, and linking design requirement with test artifacts, events, and resources, etc. as suggested by Boumen, et al. (2009), Graham (2002), and Uusitalo, Komssi, Kauppinen, & Davis (2008), it is difficult to expand RBT&E's existing limited focus and scope to objectively support the concurrent characterization of element performances-under-test as pertain to the integrated SOI design (Eigbe, et al., 2010; Ncube, 2011; Bjorkman & Zeis, 2007). Since it is not possible to physically lengthen the element verification activities with an early start prior to the SOI Implementation process (i.e., the element functions-under-test are not yet available), several existing integration-level T&E approaches and techniques have been introduced to the T&E community as partial supplements the tradition RBT&E approach and they are discussed next (Harman, et al., 2016; Uusitalo, et al., 2008).

2.3 Existing RBT&E Supplements to Forecast SOI Capability

To date, there exist only a few integrated-level test methods and approaches to supplement (not replace) the traditional RBTE approach which this study reviews in the next subsections. This study's research claim that each of these RBT&E-supplemental test approaches and methods bear its own inadequacies as pertains to produce an objective SOI-capability forecast, and none introduces a complete solution to the encountered RBT&E's data-inconsistency challenges. In response, this study's proposed SBED's methodology encapsulates each of the existing supplemental test approaches and methods' benefits and contributions; In return, a complete top-down SOI-synoptic RBT&E-supplemental verification approach to objectively forecast SOI capability is proposed by this study.

2.3.1 SE-Based T&E Framework and Methods

Several SE-oriented test frameworks and methods have been developed during the recent years intended to introduce a top-down SOI-synoptic perspective throughout the SOI Integration and Verification processes. Most of these new SE-oriented test frameworks and methods are specifically tailored toward the development of software (SW)-centric SOIs and place a heavy reliance on the availability of large SW-verification data sample size (e.g., Dou & Reascos [2015]; Zeng & Tan [2014], and Deplanques, et al. [1996]). This is a luxury typically not available to hardware (HW)-centric SOI V&V processes. Majority of reminder SE-oriented test frameworks and methods are aimed toward improving the automated generation and management of SOI-element test cases and scenarios with specific focus to the respective SOI development life cycle's resource-allocation efficiency (e.g., Boumen, et al [2009], Bjorkman, et al. [2013], and Kukula, et al.[(2012), or risk-management practices (e.g., Fiebrandt & Wilson [2008] and Bjorkman, et al. [2005]), or data-to-requirement traceability (e.g., Dahmann [2012], Helle & Schamai [2016], and Xia, et al. [2013]).

In their own regards, each of these existing SE-oriented test frameworks and methods contributes to the overall SOI development life cycle and introduces aspects of top-down SOI-synoptic verification approach to alleviate the complexities associated to overall SOI V&V processes. However, this study finds that amongst these existing SE-oriented test frameworks and methods, none specifically addresses this study's problem of objectively forecasting SOI capability during the SOI Implementation process.

2.3.2 The Early Prototype T&E Approach

Preliminary SOI-capability identification through the V&V of production-representative prototype assets is usually the most desirable option that test planners may consider supplementing the traditional RBT&E approach. During the SOI Implementation process, early prototype verification approach is typically carried out as a parallel effort (to RBT&E) that is primarily focused on the SOI user requirement(s) or few selected critical element-level requirements (Tsai, et al., 2010). For example, DoD (2017) instructs tailoring early prototype's test cases and scenarios to determine the SOI design's operational effectiveness (OE) and/or operational suitability (OS) through demonstrated compliancy towards the selected few requirements (e.g. key performance parameters, critical technical parameters, and measures of performance, etc.).

However, from a feasibility perspective, the early prototype test approach is extremely challenging to implement during the SOI Implementation process. This is because due to the independent and asynchronous nature of SOI-element development strategies, assembled early-prototype T&E assets (and thus the employment of early-prototype test approach) are usually not available (Ncube, 2011; Weiss & Schmidt, 2010; Dahmann, et al., 2010). Therefore, while the early prototype test approach introduces a schedule-compliant ability to identify SOI capability and delivers SOI mission-representative verification results that bear direct relevance towards the SOI user's operational needs (e.g., OE and OS) as reported by Eigbe, et al. (2010), a similar but more feasible top-down SOI-synoptic verification approach is needed to supplement RBT&E during the SOI Implementation process (Li, et al., 2018a).

2.3.3 MBSE and M&S Test Approaches

As previously introduced in Subsection 2.2.3 (Inadequacies of the RBT&E to Support MBSE's SOI-Capability Simulation), MBSE and other modeling and simulation (M&S) test approaches are ideal to characterizing the dynamic integrated-SOI level performances with adaptive abilities and at reduced costs (Piaszczyk, 2011; Parry, 1996; Bjorkman, et al., 2012). M&S test approaches encompass employing a methodical SOI-performance definition technique, start-to-end, through representative-interconnection and -sequencing of all applicable element (or lower-level) performance models. It then characterizes SOI performances through the aggregated and coherent simulation of these interconnected lower-level models (Dahmann, et al., 2010).

M&S is a proven test approach to inspect the SOI design, as it has been heavily utilized as so to develop and checkout unambiguous requirements during Decomposition and Definition processes (INCOSE, 2015; Bjorkman, et al., 2015; Dou & Reascos, 2015). However, once SOI development transitions to the Implementation process and beyond, M&S's actual T&E benefits towards reliably forecasting SOI capability through its aggregated coherent simulations of lower-level performance models (that are beyond repetitive and routine processes) remain undocumented and unproven (Karban, et al., 2008; Hess & Valerdi, 2010; Benjamin, et al., 1999). This is because M&S-produced performance data are often considered as untrustworthy and incomplete (INCOSE, 2015; Parry 1996). And in order to make the models into reality, heavy insistence is placed on iteratively repeating the same simulations using subsequently-collected real-life verification data as model inputs (Dahmann, et al., 2010; Nuseibeh & Easterbrook, 2000; Hynninen & Kauppinen, 2014). Thus, a continuous "model, test, model" sequential

iteration is created that may adversely obscure the Implementation process' T&E focus and resources away from its primary *Requirement-Based Performance Verification* goal (Bjorkman, et al., 2015).

In response, several model-based frameworks have been introduced specifically to imporve (not eliminate) this "model, test, model" M&S test approach. For example, Brjokman, et al. (2015) and Helle & Schamai (2016) each proposes a model-based T&E framework intended to improve the "model-to-test" process by incorporating M&S verification data to improve the development and management of actual verification test artifacts, and Zhu et al. (2016) also proposes a model-based T&E framework intended to improve the "test-to-model" process by establishing traceability between collected verification data to their respective M&S models. However, to date, there does not exist a proposed model-based T&E framework that effectively eliminates this "model, test, model" iterations by introducing an alternative that strategically integrates M&S together with actual RBT&E-collected data to objectively forecast the integrated-level capability via aggregate and coherent simulation (Bjorkman, et al., 2015).

Furthermore, and specific to the SOI Implementation process, SOI capability forecasting through MBSE's coherent simulation of respective lower-level models are encountered with the same data-domain inconsistency challenge when RBT&E verification data are utilized as respective lower-level model inputs. As previously discussed in Subsection 2.2.3 (Inadequacies of the RBT&E Approach), an SOI capability forecast, one that is representative of the SOI-mission and comprehensive to the SOI-operational environment, cannot be coherently simulated unless the available RBT&E

element verification data utilized as simulation inputs also bear the same characteristics (Winkler & von Pilgrim, 2009; Shaikh & Moore, 1997; Fiebrandt & Wilson, 2008).

Therefore, to enable the delivery of an objective SOI-capability forecast that comprehensively characterizes SOI's ability to achieve its requirements throughout its operational environment, SBED adopts M&S test approach's SOI-performance definition technique (through the representative-interconnection and -sequencing of applicable elements) and further incorporates additional DoE and RSM analytical techniques to deliver SOI-mission representative element verification data (Li, et al., 2018a).

2.3.4 Experimental Design Test Method

Experimental design test method, also commonly referred to as DoE, is a systematic performance-assessment technique intended to capture the effects from individual and combined controlled input variables, or 'factors' (Winkler & von Pilgrim, 2009). It is a mature T&E technique and has gained acceptance from many T&E communities for implementation throughout the SOI Implementation, and Integration and Verification processes (Montgomery, 2009; Myers & Montgomery, 2002; Giadrosich, 1995; Hill & Chambal, 2009).

When DoE is employed during the SOI Implementation process' element verification activities, element performance data-collection locations, or 'design points', are strategically assigned throughout the verification test's entire observation region to comprehensively capture the cause-and-effect (or causality) relationship between changes in the test observation region's parameters and changes in the observed element performances (Winkler & von Pilgrim, 2009). It objectively expands each individual

element verification's scope to a broad-spectrum vice RBT&E's static requirement specific measurement conditions (Bjorkman & Gray, 2009).

Although the DoE test method produces more-comprehensive and more-descriptive characterizations of element performances-under-test, as compared to RBT&E's scalar conditional verification data, it alone does not resolve RBT&E's data-metrological and data-domain inconsistency challenges encountered during the SOI Implementation process (Deaconu & Coleman, 2000; Bai, et al., 2001). For example, even if DoE test methodology is systemically instituted as part of the SOI verification process's overall test strategy, its methodology remains to be independently executed during each individual element verification. This implies that only the requirement-specific element variable is characterized w.r.t. the requirement-specific test observation region/operational environment (Deaconu & Coleman, 2000); therefore, all the DoE-produced element performance characterizations, albeit comprehensive to their respective allocated requirements' specified operational environments. Collectively, they remain to be another miscellany of mismatched element verification results (Li, et al., 2108a).

Furthermore, the employment of DoE test method alone during the SOI Implementation process (i.e., without the RBT&E test approach) may also adversely obscure (if not introduces additional hindrance towards) the ability to accomplish the primary *Requirement-Based Performance Verification* T&E goal. With each allocated element requirement-under-test identified as a critical function within the SOI design, its corresponding test cases and scenarios are expected to collect requirement compliance verification data as accompanied with high objective confidence (Shaikh & Moore, 1997). However, DoE's methodology of strategically assigning the finite (and often already

limited) number of data samples, or 'design points', across the test's observation region (and away from the allocated element requirement's specific verification condition) may potentially further introduces unnecessary burden towards the accomplishment of the primary *Requirement-Based Performance Verification* T&E goal (Hahn & Meeker, 1993; Hess & Valerdi, 2010).

Therefore, in the interest to resolve RBT&E's mismatch of element variable conditions challenge and to produce a detailed SOI capability forecast through the quantitative amalgamation or coherent simulation of its element performance model, SBED (in part) adopts DoE test method's comprehensive performance-modeling ability and systemically introduces the SOI operational environment as the common test observation region (or input domain) to normalize all to-be-collected element verification data.

2.3.5 SOI Mission-Relevant Assessment Analysis (Deficiency Tracking)

The SOI-element deficiency-report accounting method standardizes independent (and mismatched) element verification data w.r.t. a common and SOI mission-relevant unit of measurement (Department of the Navy [DoN], 2011). As individual element verification data are collected, and corresponding non-requirement compliant performance failures/anomalies are identified, their respective deficiency reports are created. These deficiency reports are then subjectively categorized through a generic (but standardized) SOI mission-capability impact assessment scale (DoN, 2011; USAF, 2015); and then categorically tallied to produce the overall SOI capability assessment. An example a SOI-

mission capability impact assessment scale from a DoD T&E agency is shown in Table 2-3 (DoN, 2011).

Table 2 - 3, Example of Deficiency Classification Guideline

Deficiency Description & Mission Impact	Mission Impact Category	Recommended Action
Inadequate performance to accomplish the mission(s), or high probability of system damage and personnel injury.	Part I**	Resolve prior to further T&E.
Excessive operator compensation required to accomplish the primary or alternate mission(s).	Part I*	Resolve prior to operational deployment.
	Part I	Resolve when possible.
Objectionable characteristics require significant operator compensation to attain adequate performance.	Part II	Resolve when practical.
Annoying characteristics requires minor operator compensation to attain adequate performance.	Part III	Avoid in future designs.

Note: ** Represents highest mission-impact. * Represents second highest.

The deficiency-report accounting method provides a practical means to resolve RBT&E's data-metrological inconsistency challenge. However, with each deficiency report's SOI mission-capability impact score remains to be a subjective (or judgmental) interpretation (or measurement), and of which with the primary intent to determine whether to implement a fix vice a comprehensive characterization of element performance limitations as pertain to the SOI mission (Bjorkman, et al., 2013), their tallied sum which is also subjective remains inadequate to support the objective and quantitative identification of SOI capability (Li, et al., 2108a).

As a result, SBED adopts the deficiency accounting method's element data-standardization technique (via a common SOI mission-relevant variable) but introduces the quantitative predicted "*element capability*" (or *performance-versus-requirement* threshold) response variable to standardize all calculated *element-capability* models and deliver an early SOI capability forecast.

2.3.6 Encapsulation by the Proposed SBED Test Approach

Overall, it remains a challenge for the SE and T&E communities to concurrently accomplish the new *SOI-Capability Identification* T&E goal based solely on the utilization of (mismatched) RBTE element verification data alone. This subsection has presented several existing supplemental test approaches aimed to address RBTE approach's data-inconsistency challenge. And each has its own inadequacies and shortcomings.

This study's proposed SBED test approach is another supplement to the traditional RBTE approach. SBED delivers a complete solution to the existing RBT&E's data-inconsistency challenges by encapsulating all the existing supplement test approaches and methods' following positive attributes: (a) the early prototype and M&S test approaches' method of defining systems-level performances through interconnecting and sequencing of applicable lower-level element functions; and (b) the DoE test approach's ability to comprehensively characterize element performances throughout the entire SOI's operational environment; and (c) the performance-deficiency accounting's method of introducing a common and SOI-mission relevant unit-of-measurement to standardize (otherwise mismatched) element verification results.

With the encapsulation of other existing test approaches and methods' benefits, SEBD introduces the ability to standardize and normalize individual (previously mismatched) element verification data, respectively, w.r.t. a common unit-of-measurement that is relevant to the SOI mission and a common domain that is comprehensive representative to the SOI operational environment. The SBED test approach's methodology is discussed in detail in Chapter 3.

3 The SBED Test Approach Methodology

This study researches and introduces the SBED test approach to supplement the traditional RBT&E approach. SBED resolves the SOI Implementation process' existing RBT&E data-metrological and -domain inconsistency challenges by capturing and characterizing all individual element performances data as response models w.r.t. a common SOI mission-relevant unit-of-measurement (i.e., *SOI-element capability* or the ability to satisfy allocated element requirement's specified threshold) and as influenced by a common SOI mission-representative domain (i.e., SOI's intended operational environment). SBED then either quantitatively amalgamates, or coherently simulates, the (now matching) element models to objectively produce the corresponding SOI-capability forecast. As previously illustrated in Figure 1-1, SBED is implemented through the following 5 steps:

Step 1. Describe or "scenarioize" the SOI-capability through applicable element functions slated to undergo verifications (as discussed in Subsection 3.2.1).

Step 2. Introduce the SOI's intended operational environment as the common data domain to normalize all element verification results (as discussed in Subsection 3.2.2).

Step 3. Design and execute element verification tests with DoE methodology (as discussed in Subsection 3.2.3).

Step 4. Implement RSM analytical techniques to calculate standardized *element-capability* forecasts (as discussed in Subsection 3.2.4).

Step 5. Determine the integrated SOI-capability forecast through either: (a) super positioning of the calculated *element-capability* forecasts (as discussed in

Subsection 3.2.5.1); or (b) aggregate and coherent simulation of *element-capability* forecasts (as discussed in Subsection 3.2.5.2).

3.1 Definitions and Prerequisite Assumptions

Several prerequisite definitions and assumption must be established prior to the implementation of SBED's methodology and they are discussed below.

3.1.1 Defining the 'SOI Capability'

First definition pertains to 'SOI capability'. This study defines SOI capability as both: (a) a set of assigned element functions that coherently contribute toward achieving a desired operational objective (Weiss & Schmidt, 2010); and (b) these elements' collective ability to achieve that desired operational objective throughout the SOI's intended operational environment (INCOSE, 2015; DoD, 2017; MODAF, 2005). SBED's incorporation of these two definitions takes place in SBED's Steps 1 and 2, respectively; and as mathematically expressed in Eq. 3-1 and Eq. 3-2, respectively.

3.1.2 SOI Capability-and-Element Capability Hierarchical Relationship

SBED is to be implemented under the assumption that every element requirement, that is slated for eventual performance verification during the Implementation process, is a "well crafted" and properly-reviewed deliverable from earlier SOI Decomposition and Definition processes (i.e., achievable, verifiable, and most importantly bears relevant bidirectional traceability to the parent SOI requirements capability in accordance to INCOSE [2015] and DoD [2017]). This assumption is to be applied in SBED's Step 5. It

implies that every SBED-forecasted *element-capability* failure (or inability to achieve element requirement-compliant performances under a specific operating condition as defined per (INCOSE, 2015; DoD, 2017; MODAF, 2005) can be justifiably interpreted as an eventual manifestation of a corresponding integrated SOI-level failure under the same operating condition (Winkler & von Pilgrim, 2009; Shaikh & Moore, 1997).

3.2 SBED Methodology Description

3.2.1 Step 1 Depict the SOI Capability Scenario

Implementation of SBED begins during the SOI's Decomposition and Definition processes when the SOI capability, start-to-end, is "scenarioized" through a sequencing set of applicable lower-level elements that are slated to undergo eventual verifications. Composition of this SOI-capability scenario diagram is conceptually described in Eq. 3-1 as follows: Given an SOI-level capability, $\boldsymbol{F}$, there exists a set, $\boldsymbol{G}$, of applicable and predefined element requirements, $\boldsymbol{f}_i$s, *where* $i = 1,2,3, \ldots, m$, that can be sequentially interconnected to depict $\boldsymbol{F}$ from its start to end.

$$\boldsymbol{F} \stackrel{\text{def}}{=} \mathbf{G}: \{\forall \boldsymbol{f}_i \rightarrow F\}, \; i = 1, 2, 3, \ldots, m; \qquad \text{(Eq. 3-1)}$$

SBED's SOI-capability definition methodology through hierarchical requirement decomposition is an existing and mature Requirement Engineering (RE) technique that is widely employed throughout the SOI Decomposition and Definition processes (Winkler & von Pilgrim 2009; Graham 2002; and Heitmeyer, Labaw, & Kiskis, 1995). It is a component of many existing RE-based T&E frameworks (e.g., DoD, 2011; Sheehan, et al.,

2004; Davis & Bigelow, 2003). Its proven benefits, towards the improvement of both SOI-design verification test artifacts' design quality (de Santiago Junior & Vijaykumar, 2012; Bjorkman, et al., 2012) and test-resources allocation (Flammini, et al., 2009), are all well documented. Thus, such an SOI-capability scenario diagram may already be available at the onset of SBED's implementation during the Implementation process as derivations of prior SOI Decomposition and Definition process' outputs. However, if such an SOI-capability scenario diagram is not readily available, it can be developed by test planners through interpretations of SOI's applicable requirement verification and traceability matrices and/or other corresponding Technical Management processes' outputs (e.g., work breakdown structure, hierarchical architecture frameworks as demonstrated from works in Liu, et al. [2012], Uuterkalmsteiner, et al. [2014] and Eppinger, Joglekar, & Gorschek [2014]).

As a visual aid, a SOI scenario diagram depicting an arbitrary to-be-forecasted SOI capability is shown in Figure 3-1. It consists of four (hypothetical) interacting and sequenced lower-level elements, $\boldsymbol{f_1} - \boldsymbol{f_4}$, and is presented as a MBSE SysML requirement and behavior-activity diagrams.

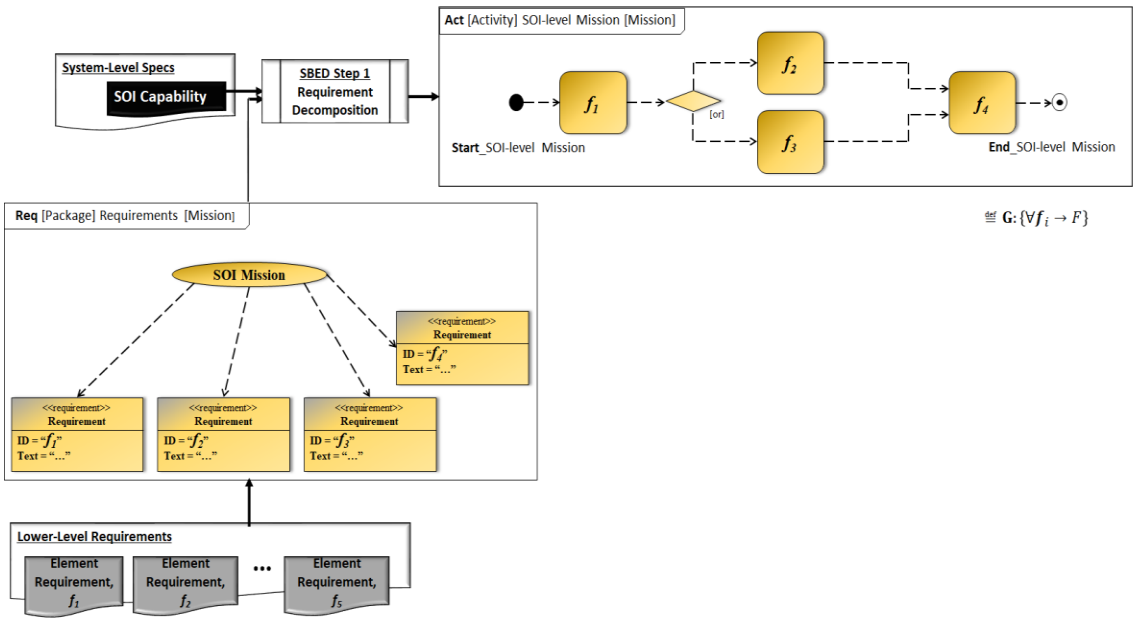


Figure 3 - 1, An Arbitrary SOI-Capability Scenario Diagram

3.2.2 Step 2 Introduce SOI Operational Environment as the Common Domain

SBED's first aim is to identify and introduce a common SOI mission-representative domain set to normalize all (previously mismatched) RBT&E element verification data. Knowing that decision makers and the T&E community both desire to comprehensively assess the SOI design throughout realistic conditions representative of its operational environment (i.e., the *SOI-Capability Identification* T&E goal as stated in DoD [2017], Bjorkman & Gray [2009], and Bai, et al. [2001]), SBED utilizes the SOI operational environment, ***R***, and its parameters, $(\xi_1, \xi_2, \ldots, \xi_k)$, as its common domain set to normalize all the yet-to-be-collected element verification data. (Standardizing them is to be accomplished later in SBED's Step 4.) The introduction of this common data-domain set is conceptually expressed as follows in Eq. 3-2: If an SOI performance is subjected for

assessment throughout its entire operational environment (i.e., SOI capability), then verifications and assessments of all its applicable child elements are also subjected.

$$\boldsymbol{F}(\boldsymbol{R}) \stackrel{\text{def}}{\equiv} \boldsymbol{G}(\boldsymbol{R}) = \boldsymbol{G}(\xi_1, \xi_2, \ldots, \xi_k) = \{\, \forall \boldsymbol{f}_i(\xi_1, \xi_2, \ldots, \xi_k) \rightarrow \boldsymbol{F}(\xi_1, \xi_2, \ldots, \xi_k) \,\},$$

where i = 1, 2, …, m; and $\boldsymbol{R} = \{\xi_{\mathbf{1}_{min}},\ \xi_{\mathbf{1}_{max}}\},\ \{\xi_{\mathbf{2}_{min}},\ \xi_{\mathbf{2}_{max}}\},\ \ldots, \{\xi_{\boldsymbol{k}_{min}},\ \xi_{\boldsymbol{k}_{max}}\}.$

(Eq. 3-2)

The parameters of $\boldsymbol{R}$, ξ_js, can be a mixture of both a system's environmental factors and conditional factors internal and external to the SOI boundary. These parameters, along with their boundaries, are to be either identified by test planners and subject matter experts or directly extracted from SOI's Stakeholder Needs and Requirements Definition process's outputs (e.g., life cycle concept documents, OpsCon descriptions, etc. as demonstrated in works from Liu et al. [2012]). For example, common data-domain set of an airborne platform's flying-quality capability can be defined w.r.t. both its operational environment factors (e.g., wind, temperature, or altitude) and its conditional factors (e.g., velocity, attitude, or aircrew experiences) per DoN (2011). For a maritime underwater vessel's navigation capability, its common data-domain set can be defined by its environmental factors (e.g., current speed and direction, depth, or ambient luminance) and conditional factors (e.g., payload weight, or mission duration) per DoN (2009).

At this juncture, the identified SOI operational environment, $\boldsymbol{R}$, and its parameters, ξ_js, are to be incorporated as the common controlled input factors for all the to-be-executed element verifications' DoE test methods and to determine the causality relationships between changes in the SOI operational environment and each of the element-under-test's demonstrated performances.

3.2.3 Step 3 Design and Execute Element Verifications with DoE Methods

The capturing and collection of element-under-tests' performances and verification data, all as normalized cause-and-effect relations w.r.t. the integrated SOI's operational environment and its parameters, $\boldsymbol{R}(\xi_j)$, *for* $j = 1, \ldots, k$, is achieved through the employment of appropriate DoE test designs in which ξ_js are incorporated as the common controlled DoE input factors for every element verification test. During the respective element verification test-planning efforts, the aim of each DoE test design is to strategically assign performance-data collection locations (or DoE 'design points') throughout $\boldsymbol{R}(\xi_j)$ (or the 'test observation region); and if possible, with a statistically adequate amount of sample size at each DoE design points (i.e., replicates).

In the interest of test-resource allocation management (i.e., to efficiently capture the "adequate" amount of causality relationship without executing unnecessary additional DoE design points), SBED and other existing DoE-related texts recommend each independent element verification activity to begin with an initial screening DoE design to conclusively remove insignificant (or inconsequential) factors and effects from any subsequent detailed DoE methods and RSM causality analysis (Montgomery [2009], Bateson and Mendel [2013], and Wu & Hamada [2009]). If, and only if, the initial screening DoE design's captured data is determined to be insufficient as indicated by the corresponding Analysis of Variance (ANOVA) results and its inherent unfavorable Lack-of-Fit sum of squares (Montgomery, 2009), SBED then recommends each independent

element verification activity to further implement 2nd-order sequential DoE designs (Ott & Longnecker [2015] and Neter, Wasserman, & Kutner [1996]).

As previously mentioned in this dissertation's Subsection 1.6 (Scope and Limitations), it is not this study's or SBED's intent to teach specific DoE designs (screening or sequential) to be employed as part of its methodology during the SOI Implementation process. Instead, the actual selections of appropriate DoE designs specific to each SOI's element verification test should remain as unique case-by-case decisions that are based on both element verification's test-resource availabilities and its intrinsic *Requirement-Based Performance Verification* objectives. The latter is to be further discussed in Subsection 3.2.3.3 (Alignment of the Primary Element Verification's Goals and Objectives). From SBED's perspective, as long as the following two considerations are applied during each element verification's DoE design selection process, SBED's methodology can continue:

1. Inclusion of the previously-defined common input domain, $R \stackrel{\text{def}}{=} \{\xi_1, \xi_2, \dots, \xi_k\}$, as its primary controlled input factors (until conclusively proven to bear insignificant effects towards the element performance-under-test); and
2. (To the maximum extent within the available test resources constraints.) Achieve of a minimum of Resolution-V design to ensure that all significant primary input factors' main and binary-interaction effects are not mutually aliased nor confounded (Montgomery, 2009). This consideration process is to be further discussed in Subsection 3.2.3.2 (Sequential DoE Test Methods).

Considerations toward the selection and execution of SBED's initial screening and subsequent sequential DoE designs are respectively discussed in the following two subsections of this dissertation. Furthermore, additional considerations for selecting each element verification's sequential DoE designs to further enhance the accomplishment of its intrinsic *Requirement-Based Performance Verification* T&E goal is also discussed.

3.2.3.1 Initial Screening DoE Test Methods

Each element verification activity should begin with a preliminary screening DoE design that utilizes near twenty-five percent of the available T&E resources (Montgomery, 2009). The primary intent of each element verification's initial DoE design, in the interest of test-resource allocation management, is to identify the parameters within SBED's assigned common controlled input factors (or Step 2's ξ_js) that bear significant effects to the element performance-under-test at the 1st-order level , or standalone main effects with or without binary-interactions as described by Montgomery (2009) and Gunter (2014). This is typically accomplished through the execution of a two-level factorial (full or fractional) screening DoE designs (Box, Hunter, & Hunter, 2005; Box & Draper, 2007) and (test-resources permitting) with a design resolution of up to either Resolution-III (i.e., all main factor effects are not aliased or confounded to each other) or a Resolution-IV design (i.e., all main factor effects are not aliased or confounded to neither each other nor to any binary-interactions (Montgomery, 2009; Box & Draper, 2007).

The fundamental benefit of executing a prerequisite screening DoE design prior to sequential DoE designs is the reduction in the number of input factors that make designing the sequential DoE test(s) more feasible and manageable. It also introduces two additional

practical benefits. The first benefit pertains to test-resource management. With the majority of sequential DoE (2^{nd}-order) experimental designs embeds a foundational 1^{st}-order factorial design, having the initial screening DoE design and its corresponding 1^{st}-order data already executed and captured allow test planners to implement the sequential 2^{nd}-order DoE designs without repeating the 1^{st}-order test trails (Montgomery, 2009). The second benefit is the ability to develop a more objective and conclusive 1^{st}-order model with the identified insignificant factors removed through iterative ANOVA analysis and without the need of repeating the 1^{st}-order test trials (Montgomery, 2009; Neter et al., 1985). The Projection Property states that factorial designs can be projected into stronger (larger) designs in the subset of significant factors (without any actual testing) and create a more sufficient 1^{st}-order regression model, assuming that the removed factors are inert (Cheng, 2006). Upon the completion of each element verification's initial screening DoE test and based on the collective interpretation of the following DoE principles listed below, SBED can conclusively remove the remaining non-significant parameters or factors (Step 2's ξ_js) from the respective element-under-test's subsequent sequential DoE test designs (if needed) to minimize the quantity of unnecessary DoE design points (Montgomery, 2009; Gunter, 2014). The DoE principles addressed are:

- **The Effect-Sparsity Principle** states that the number of relatively important effects is small as described in Lenth (1989) and Box & Myer (1986); and
- **The Effect Heredity Principle** which states factors which produce insignificant main effects usually show no significant interactions (i.e., for a multi-factor interaction to be significant, at least one of its parent factors should

be significant) as described in Chipman (1996), Nelder (1998), Bateson & Mendel (2013) and Wu & Hamada (2009); and

- **The Effect Hierarchy Principle** states that lower-order effects are more important (or significant and 'interesting') than higher order effects as described in Montgomery, Myers, Carter, & Vining (2005), Chipman (1996), and Berger, Knowlton, & Abrahams, (1996);

3.2.3.2 Sequential DoE Test Methods

There exists a likelihood that the initial screening DoE design' outputted 1st-order cause-and-effect relationship descriptions, as obtained through subsequent ANOVA analysis (Ott & Longnecker, 2015), is in need of further refinement or is determined to be insufficient either through the objective lack-of-fit testing that partitions residual sum of squares (Montgomery, 2009; Myers & Montgomery, 2002) or the subjective subject-matter-expert interpretation of the empirical model). This requires subsequent sequential DoE designs to be employed to develop respective element performance-under-test's 2nd-order regression models and to capture the remaining (significant) parameters or factors' quadratic effects.

There are numerous proven and documented 2nd-order sequential DoE (or RSM) designs available. The most straight forward choice for a sequential DoE design is the three-level factorial designs (Montgomery 2002, Khuri & Cornel, 1987). The three-level factorial designs, full or fractional (also known as the Box Behnken Design per Box & Behnken [1960]), maximizes DoE's Projection Property advantage from previous screening designs to capture 2nd-order or quadratic factor-effect estimate. It is the most

efficient sequential design because the least number of additional test trials is required, especially when the identified SOI operational enviornment, $\boldsymbol{R}$, has only a few parameters (i.e., $k \leq 3$). However, the three-level factorial designs are often difficult to implement because it requires design points (or data-collection locations) at the most extreme of several (if not all) operational and conditional parameter values (i.e., unreasonable operating conditions). For example, the airborne platform's performance under maximum wind, most extreme temperature, and highest altitude while conducting the most demanding maneuvers at maximum cruising velocity. These extreme test conditions may not always be feasible (or safe) during the SOI Implementation process' initial element verification activities. As a result, test planners often elect to select the most common and practical choice as part of their sequential DoE design. This leads to the employment of variations of the Central Composite Designs (CCD) as discussed in Myers & Montgomery (2002) and Khuri & Cornell (1987). The CCD designs capture 2^{nd}-order factor-effect estimates through more reasonable and more feasible design points where only one parameter is set to its extreme value at any given time. The remainder parameters are set to neutral. Variations of the CCD design are distinguished amongst the magnitude of the extreme parameter's value, typically either at the defined parameter boundary per Eq. 3-2 or as a function of the number of identified SOI operational enviornment, k, per Myers (2002). This flexibility amongst CCD design introduces a test-planning advantage which allows the selected DoE design to concurrently accommodate the intrinsic nature of each element verification's *Requirement-Based Performance Verification* T&E goal. This advantage, also known as DoE design optimality, is further discussed in the Subsection 3.2.3.3.

When specific element verification test's available resource limitations and constraints prohibit the implementation of three-level factorial (full or fractional) or CCD sequential designs, numerous saturated and less-practiced 2^{nd}-order designs are also available to capture 2^{nd}-order factor-effect estimates. Some examples include the Plackett-Burman (screening) Design per Plackett & Burman (1946), the Hoke Design per Hoke (1974), the Koshal Design per Koshal (1933), the Notz Design as introduced in Myers & Montgomery (2002), and other hybrid designs are mentioned in Roquemore (1976), Khuri & Cornell (1996), and Myers & Montgomery (2002). Implementation of any saturated DoE designs requires a smaller number of design points (or test trials and resources) than the three-level factorial or CCD; however, this advantage is accompanied with the cost that a number of factor-effects cannot be quantitatively estimated and must be intentionally removed from subsequent RSM analysis and modeling (i.e., 'saturated').

In all, while each of these sequential DoE designs incorporates unique design point and test trail emplacement locations, they are all capable of outputting similar 2^{nd}-order cause-and-effect relation descriptions between changes in the SOI operational environment and each of the element-under-test's demonstrated performances. Thus, to conclusively capture the 2^{nd}-order cause-and-effect regression models, SBED recommends that a minimum of Resolution-V DoE design is selected (i.e., both main and binary-interaction effects are not mutually aliased or confounded to each other per Montgomery, 2009 and Box & Draper, 2007). However, each element verification's independent selection of its 2^{nd}-order DoE designs is ultimately based upon the following two influences (in order of priority): (a) the extent of allocated test resources' ability to implement each DoE design's required trail sizes and locations; and (b) the DoE design's ability (or 'optimality') to fulfill

the intrinsic nature of each element verification's *Requirement-Based Performance Verification* T&E goal, as further explained in the Subsection 3.2.3.3.

3.2.3.3 Alignment to the Primary Element Verification's Goals and Objectives

To adhere to each element verification test's inherent *Requirement-Based Performance Verification* goal that is w.r.t. their respective allocated requirement's threshold and verification condition, each SBED-employed DoE design must also include its allocated element requirement-under-test's specified verification condition(s) as part of its actual data-collection location(s) as DoE design points. If the respective element requirement's specified verification condition(s) is defined by additional parameter(s) that are not part of the SBED Step 2's predefined SOI operational environment, $\boldsymbol{R} \stackrel{\text{def}}{=} \{\xi_1, \xi_2, \ldots, \xi_k\}$, SBED recommends that the element verification test's respective DoE design(s) consider including these specific parameter(s) as its additional controlled DoE input factors vice just as inert constant scalar inputs.

Additionally, SBED further encourages considering the unique, intrinsic nature and expectation of each element performance-under-test's *Requirement-Based Performance Verification* goal as part of its sequential DoE design selection process. Typically, element performance verification data are expected to conclusively describe one of the following: (a) the satisfaction of the implemented element performance-under-test w.r.t. the requirement's specified verification condition(s); or (b) conversely the suitability of element requirement-introduced verification condition(s) w.r.t. the element functionality-under-test. While these two *Requirement-Based Performance Verification* expectations do

not warrant differences in the reporting structure of traditional RBT&E verification data (i.e., continue to conclusively demonstrate y_1-*mph* speed under $x_1{}^oC$ verification condition, y_2-*tons* payload at x_2-*m* altitude, y_3-*Watts* sustain output power after x_3-*hours* of operation, etc.). However, these difference in expectations can be specifically and distinctly addressed based on the flexibilities within 2nd-order DoE design characteristics (i.e., design optimality efficiency). To ensure the conclusive confirmation of the implemented element performance-under-test's satisfaction w.r.t. the requirement's specified verification condition(s), the key focus of a sequential DoE design is to generate fitted regression models that produce lower prediction variances at those location(s) of significance. This can be achieved by selecting DoE designs with high G-optimal (or IV-optimal) efficiency as discussed in Box & Hunter (1957) or IV-optimal efficiency as discussed in Hardin & Sloane (1993). Conversely, to ensure the conclusive confirmation of the allocated requirement-specified condition(s)'s suitability w.r.t. the element function-under-test, the key focus of respective DoE design shifts to generating fitted regression models with the least average variance for either: (a) all included regression coefficients by selecting DoE designs with A-optimal efficiency as discussed in Chernoff (1953), or (b) a preselected group of regression coefficients by selecting DoE designs with high D-optimal efficiency as discussed in Kiefer (1959) and Goos & Jones (2011). The detailed mathematical descriptions as pertain to quantifying a DoE design's G-, IV-, D-, and A-optimality efficiencies are presented in the works of Myers & Montgomery (2002), Kuhri & Cornell (1987), and Oyejola & Nwanya (2015), but considered to be outside the scope of this dissertation's content.

3.2.3.4 Capture Element Performance and Verification Data

Through the employment of appropriately selected DoE designs, SBED captures and collects each element-under-test's performance and verification data as normalized causality relationships w.r.t. the integrated SOI's operational environment and its parameters, $\boldsymbol{R}(\xi_j)$, *for* $j = 1, \ldots, k$. Table 3-1 presents an example of an element performance and verification data set through DoE test methods. It includes an element verification data set consists of $n > k$ test conditions (or 'design points') and $N \geq n$ trials (i.e., potential replications at certain test conditions depending on the selected DoE design).

After DoE test methods have been employed during every element verification test, the cause-and-effect relationship between changes throughout the entire $R(\xi j s)$ and changes to the demonstrated element performance that are embedded within the respective raw verification data. These causality relationships are to be further identified, quantified, and interpolated through subsequent RSM analysis during SBED's Step 4.

Table 3 - 1, Example of Raw Design of Experiment Performance Verification Data

Collected Test Results	**Input Variables (ξ_js)**			
y_{f_i}	ξ_1	ξ_2	…	ξ_k
y_1	$\xi_{1,1}$	$\xi_{1,2}$	…	$\xi_{1,k}$
y_2	$\xi_{2,1}$	$\xi_{2,2}$	…	$\xi_{2,k}$
.	.	.		.
.	.	.		.
.	.	.		.
$y_{N_{f_i}}$	$\xi_{N_{f_i},1}$	$\xi_{N_{f_i},2}$	…	$\xi_{N_{f_i},k}$

3.2.4 Step 4 Calculate Standardized *Element-Capability* Forecast Models

Once each element verification's raw performance verification data, $\boldsymbol{y_{f_i}}$, is collected (as shown in Table 3-1), the respective element linear-regression mean-response model, $\boldsymbol{\hat{y}_{f_i}}$ or $\boldsymbol{\hat{f}_i(R)}$, and prediction confidence-interval model, $\boldsymbol{\hat{f}_{i_variability}(R)}$, can be calculated using RSM statistical techniques (Myers & Montgomery, 2002; Box & Draper, 2007). Subsequently, SBED develops a corresponding *element-capability* model, $\boldsymbol{\hat{f}_{i_capability}(R)}$, through statistical comparisons between combination of $\boldsymbol{\hat{f}_i(R)}$ and $\boldsymbol{\hat{f}_{i_variability}(R)}$ and the corresponding element requirement thresholds, T_{f_i}. The derivation methods of these regression models are detailed in the following subsections.

3.2.4.1 RSM Analysis Setups

3.2.4.1.1 Conversion to Coded ζ_j Input-Variable Values

Before analyzing each element performance and verification data to produce the respective regression models, all of $\boldsymbol{R}$'s parameters, ξ_js as previously identified in Step 2 (Eq. 3-2), are individually normalized into coded variable valued between {-1, 1} as instructed in Montgomery (2009) and Box & Draper (2007). Eq. 3-3 defines the mathematical conversion between the presentation of $\boldsymbol{R}$'s parameters in coded forms. The coded variables of ξ_js, x_js, are expressed as dimensionless variables with means, μ, equal to zero and with identical standard deviations, σs. Table 3-1 further illustrates this relationship by expanding Table 3-1's raw DoE verification performance data and its corresponding coded ξ_js equivalents, x_j.

$$x_j = \frac{\xi_j - [max(\xi_j) + min(\xi_j)]/2}{[max(\xi_j) - min(\xi_j)]/2},$$

where max(ξ_j) and min(ξ_j) are defined as the boundary of each parameter within ***R*** *from Eq. 3-2.* (Eq. 3-3)

Table 3 - 2, DoE Design-Point Conversions

Input Variables, ξ_j				**Input Variables, x_j (coded)**			
ξ_1	ξ_2	…	ξ_k	x_1	x_2	…	x_k
ξ_{11}	ξ_{12}	…	ξ_{1k}	x_{11}	x_{12}	…	x_{1k}
ξ_{21}	ξ_{22}	…	ξ_{2k}	x_{21}	x_{22}	…	x_{2k}
•	•		•	•	•		•
•	•		•	•	•		•
•	•		•	•	•		•
ξ_{N1}	ξ_{N2}	…	ξ_{Nk}	x_{N1}	x_{N2}	…	x_{Nk}

For the remainder of SBED, all subsequent RSM calculations are to be performed w.r.t these corresponding coded variables of x_js.

3.2.4.1.2 Partitioning of ***R*** *for Point-Estimate Calculations*

With the SOI operational environment, ***R***, defined by a set of predetermined parameters, ξ_js or x_js *for j = 1, …, k*, it is to be visualized as an *k*-dimensional space that is theoretically partitioned into $x_R{}^k$ number of identically-sized *k*-dimensional subspaces, R_{x_R}s. This is to support subsequent regression model's point-estimate calculations. The value of x_R defines the presentation resolution of RSM-formulated regression models and how many corresponding point estimates within the models are to be derived.

It should be noted that SBED's selection of its x_R value is appropriately based on: (a) the desired model-presentation granularity of each element verification data as specified within the SOI and/or element requirement verification tolerances (e.g., element A must

demonstrate y_1-mph speed under x_1°C $\pm$ % verification condition); and (b) the available computation resources' capability (i.e., the allotted quantity of automated repetitive point-estimate calculations). For example, this dissertation's data-analysis example selects a x_R value of 0.05 (coded) within a (k=4)-dimensional $\boldsymbol{R}$ that is based on the computational capability of generating $(2/0.05)^{k=4} > 2.8 \cdot 10^6$ individual point estimates (per each computed regression model, per element function) via MATLAB. Furthermore, it should be noted that the value of x_R can also be modified appropriately throughout SBED Step 4's analysis effort (if needed) without the need of altering the raw DoE data, $\boldsymbol{y_{f_i}}$, or repeating any actual and already-performed element verification activities.

An example of a (k=2)-dimensional $\boldsymbol{R}$ visualized as an XY plane and partitioned in $(x_R)^k$ number of subspaces, R_{x_R}s, is illustrated in Figure 3-2. It is also mathematically expressed in Eq. 3-4. Similarly, any ($k \geq 3$)-dimensional $\boldsymbol{R}$ spaces are to be visualized as 3-dimensional XYZ spaces with the remaining (k-3) ξ_js presented as x_R number of incremental constants. A (k=3)-dimensional $\boldsymbol{R}$ space is mathematically expressed in Eq. 3-5.

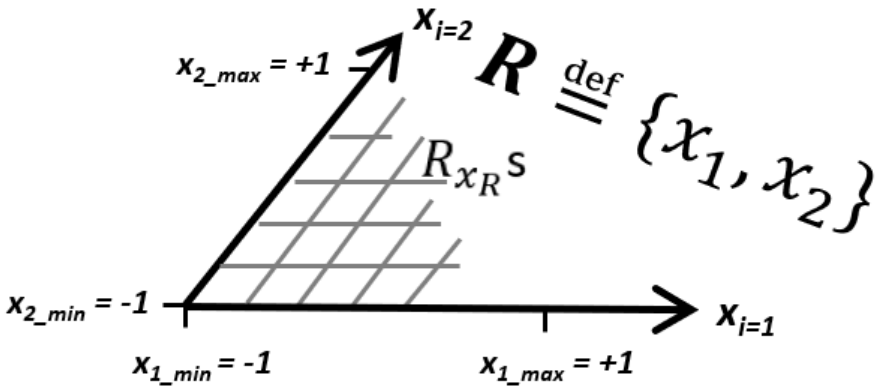


Figure 3 - 2, A 2-Dimensional SOI Operational Environment Domain

$\boldsymbol{R} \overset{\text{def}}{\equiv} \{R_{l,m}\}$, *where l, m = 1, 2, …,* x_R

$$= \begin{bmatrix} R_{1,1} & \cdots & R_{x_R,1} \\ \vdots & \ddots & \vdots \\ R_{1,x_R} & \cdots & R_{x_R,x_R} \end{bmatrix} \quad \text{(Eq. 3-4)}$$

$\boldsymbol{R} \overset{\text{def}}{\equiv} \{R_{i,j,l}\}$, *where i, j, l = 1, 2, …,* x_R; *and*

$\boldsymbol{R}$=

$$\left\{ \begin{bmatrix} R_{1,1,1} & \cdots & R_{x_R,1,1} \\ \vdots & \ddots & \vdots \\ R_{1,x_R,1} & \cdots & R_{x_R,x_R,1} \end{bmatrix}, \begin{bmatrix} R_{1,1,2} & \cdots & R_{x_R,1,2} \\ \vdots & \ddots & \vdots \\ R_{1,x_R,2} & \cdots & R_{x_R,x_R,2} \end{bmatrix}, \ldots, \begin{bmatrix} R_{1,1,x_R} & \cdots & R_{x_R,1,x_R} \\ \vdots & \ddots & \vdots \\ R_{1,x_R,x_R} & \cdots & R_{x_R,x_R,x_R} \end{bmatrix} \right\}$$

(Eq. 3-5)

3.2.4.2 Generating Element Performance Mean-Response Models

3.2.4.2.1 Calculate Elements' Mean Response

Depending on the true organic behavior of each element's performance and the extensiveness of the DoE design(s) employed to capture that behavior (i.e., the quantity and emplacement locations of 'design points' within the test observation region, *n*, and the quantity of total trails executed, *N*), the elements' respective linear-regression mean-response models, $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$s, are calculated and typically expressed as first- or second-order polynomials shown in Eq. 3-6. Both the Effect-Sparsity Principle and the Effect-Hierarchy Principle define that there is no need to analyze the SOI operational environment parameters' main- and interaction- effects at higher orders (Montgomery, 2009; Box & Meyer, 1986). Hence the importance of achieving Resolution-V DoE designs to the maximum extent possible, as discussed in Subsection 3.2.3.2 of this dissertation (Sequential DoE Test Methods).

$$\hat{\boldsymbol{f}}_i(\boldsymbol{R})_{first} = \beta_0 + \sum_j^k \beta_{f_i,j} \cdot \zeta_j + \sum_{j,l,j\neq l}^k \beta_{f_i,j,l} \cdot \zeta_j \cdot \zeta_l + \epsilon_{f_i}$$

$$\hat{\boldsymbol{f}}_i(\boldsymbol{R})_{second} = \hat{\boldsymbol{f}}_i(\boldsymbol{R})_{first} + \sum_j^k \beta_{f_i,jj} \cdot {\zeta_j}^2 + \epsilon_{f_i} \quad \text{(Eq. 3-6)}$$

The method of least squares is employed to calculate each element's respective fitted regression model (Montgomery, 2009; Lay, 2003). It sets all the least-square normal equations to zero so that the sum-of-squares of error, ${\epsilon_{f_i}}^2$, is minimized, and then solves for each of the regression model's regression coefficients, β_{f_i}s. The solutions to the least-square normal equations are β_{f_i}s' least-square estimators $\hat{b}_{f_i}$s, or $\widehat{\boldsymbol{b}}_{\boldsymbol{f_i}}$, (Myers & Montgomery, 2002; Box & Draper, 2007; Hogg & Craig, 2005); and are rudimentarily derived if expressed in matrix notations (as shown in Eq. 3-7) where in general: $\boldsymbol{y}_{\boldsymbol{f_i}} = [y_1, y_2, \ldots, y_{N_{f_i}}]$' is a (*N* by 1)-vector of raw collected element performance and verification data (as previously shown in Table 3-2), and $\boldsymbol{X}_{\boldsymbol{f_i}} = \begin{bmatrix} 1 & \mathrm{x}_{11} & \mathrm{x}_{12} & \ldots & \mathrm{x}_{1k} \\ 1 & \mathrm{x}_{21} & \mathrm{x}_{22} & \ldots & \mathrm{x}_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \mathrm{x}_{N1} & \mathrm{x}_{N2} & \cdots & \mathrm{x}_{Nk} \end{bmatrix}$ is an (*N* by *k*+1)-Design Matrix portraying the levels of the input variables used as part of element verification's DoE designs (also as shown in Table 3-2), $\boldsymbol{\beta}_{\boldsymbol{f_i}} = [\beta_0, \beta_1, \ldots, \beta_k]$' is a (*k*+1 by 1)-vector of the regression coefficients (to be set as unknown variables), and $\boldsymbol{\varepsilon}_{\boldsymbol{f_i}} = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N]$' is an (*N* by 1)-vector of random errors (to set to be minimized). If the least-square normal equations (from Eq. 3-7) are all set to zero. Then the least-square estimators of $\boldsymbol{\beta}_{\boldsymbol{f_i}}$ can be algebraically rearranged and solved as their estimates, $\widehat{\boldsymbol{b}}_{\boldsymbol{f_i}} = [\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_k]$ or $\hat{b}_{f_i}$s, with the application of Eq. 3-8 (Myers & Montgomery, 2002; Hogg & Craig, 2005).

$$\boldsymbol{y}_{\boldsymbol{f_i}} = \boldsymbol{X}_{\boldsymbol{f_i}} \cdot \boldsymbol{\beta}_{\boldsymbol{f_i}} + \boldsymbol{\varepsilon}_{\boldsymbol{f_i}} \quad \text{(Eq. 3-7)}$$

$$\boldsymbol{\beta}_{f_i} \stackrel{\text{def}}{=} \widehat{\boldsymbol{b}}_{f_i} = (\boldsymbol{X}_{f_i}{}'\boldsymbol{X}_{f_i})^{-1} \cdot \boldsymbol{X}_{f_i}{}'\boldsymbol{y}_{f_i} \qquad \text{(Eq. 3-8)}$$

Consequentially and explained in Eq. 3-9, each element mean-response model, $\widehat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$, and its β_{f_i}s, previously expressed as polynomials in Eq. 3-6, can now be alternatively rewritten and expressed as numerical array of $\widehat{\boldsymbol{b}}_{\boldsymbol{f_i}}$ and without the need to account for the unknown observed errors, ε_is, (Myers & Montgomery, 2002).

$$\boldsymbol{y}_{\boldsymbol{f_i}} = \mathbf{X}_{\boldsymbol{f_i}} \cdot \boldsymbol{\beta}_{\boldsymbol{f_i}} + \boldsymbol{\varepsilon} \rightarrow \quad \widehat{\boldsymbol{y}}_{\boldsymbol{f_i}} = \widehat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R}) = \mathbf{X}_{\boldsymbol{f_i}} \cdot \widehat{\boldsymbol{b}}_{\boldsymbol{f_i}} \qquad \text{(Eq. 3-9)}$$

This method of least squares is to be implemented to calculate every elements' respective linear-regression mean-response models, $\widehat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$s. It is also reapplied to determine every SOI expected element's performance w.r.t. the entire SOI operational environment. This is achieved by applying the calculated $\widehat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$s to derive individual corresponding point estimates, $\hat{f}_i(R_{x_R})$s.

3.2.4.2.2 Derive Elements' Prediction Point Estimates

SBED repetitively substitute each element's calculated mean-response model, $\widehat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$, to derive $(x_R)^k$ number of scalar mean-response point estimates, $\hat{f}_i(R_{x_R})$s. One at every R_{x_R} subspace throughout $\boldsymbol{R}$. As explained in Eq. 3-10, each scalar mean-response point estimate, $\hat{f}_i(R_{x_R})$, is derived by first redefining its respective R_{x_R} subspace's operating condition as a moment matrix, $\boldsymbol{x_{(m)}}$ or $\boldsymbol{x_0}$, w.r.t. the element verification's DoE

Design Matrix, $\boldsymbol{X_{f_i}}$ (from Eq. 3-7 to Eq. 3-9), and then calculates Eq. 3-11 (as derived from Eq. 3-9) with $\boldsymbol{x_0}$ substituted as $\boldsymbol{X_{f_i}}$ (Myers & Montgomery, 2002, p. 318). This substitutive-derivation of a $\hat{f}_i(R_{x_R})$ is further illustrated in Figure 3-3.

$$R_{x_R} \overset{\text{def}}{=} \mathbf{x}_0' = [1, \zeta_1^*, \ldots, \zeta_k^*, \zeta_1^* \cdot \zeta_2^*, \ldots, \zeta_{k-1}^* \cdot \zeta_k^*, (\zeta_1^*)^2, \ldots, (\zeta_k^*)^2], \textit{ where } \mathbf{x_0} \in \boldsymbol{R} \quad \text{(Eq. 3-10)}$$

$$\hat{f}_i(\mathbf{x_0}) = \hat{f}_i(R_{x_R}) = \mathbf{x}_0' \cdot \hat{\boldsymbol{b}}_{\boldsymbol{f_i}} \quad \text{(Eq. 3-11)}$$

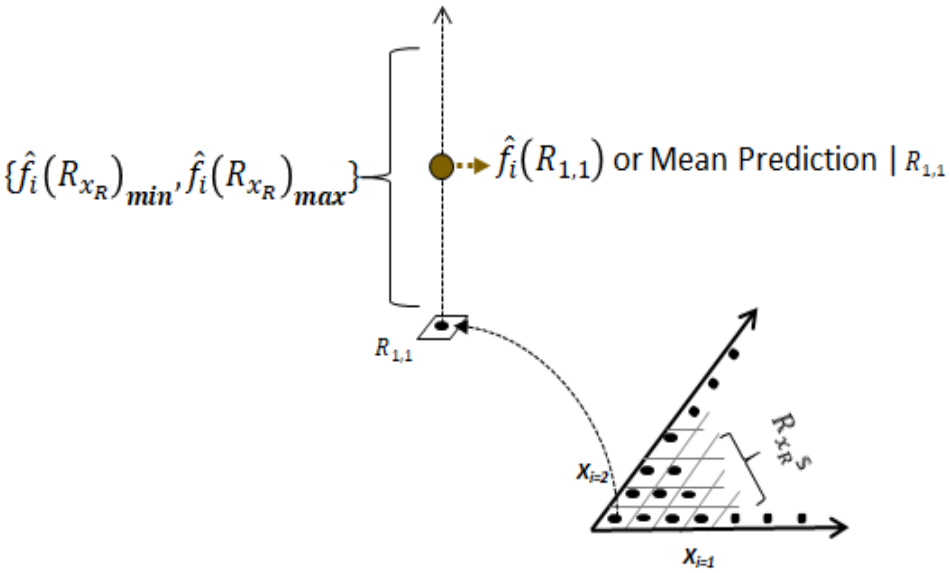


Figure 3 - 3, Calculated Mean-Response Point Estimate

It is worthwhile to caution any extrapolation (vice interpolation) of any $\hat{f}_i(R_{x_R})$s beyond the actual boundaries of $\boldsymbol{R}$ as defined in Eq. 3-2. Extrapolation of those $\hat{f}_i(R_{x_R})$s, in theory, only adequately described the cause-and-effect relations between element

performances and changes in $\boldsymbol{R}(\xi_j$s) as demonstrated through the captured raw performance data, $\boldsymbol{y_{f_i}}$. They may not reliably do so outside the DoE test observation region where different (and potentially not yet captured) causality may exist (Montgomery, 2009; Khuri & Cornell, 1996). Thus, this further emphasizes the importance of including the $\xi_{i_{maximum}}$ and $\xi_{i_{minimum}}$ of as the data-domain set for every SOI operational-environment parameters during SBED's Step 2.

3.2.4.2.3 Construct Element Mean-Response Models

Once all the corresponding element mean-response point estimates, $\hat{f}_i(R_{x_R})$s, are individually derived for all R_{x_R} subspaces, they are to be consolidated into the *k*-dimensional numerical array of $\boldsymbol{R}$ to construct the respective element's mean-response model, $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$. Mathematical expressions of constructed $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$s within a (*k*=2)-dimensional numerical array of $\boldsymbol{R}$ from Eq. 3-4 is shown Eq. 3-12 and illustrated in Fig 3-4 below as a collective construct of Figure 3-3.

$$\hat{\boldsymbol{f}}_i(\boldsymbol{R}) = \begin{bmatrix} \hat{f}_i(R_{1,1}), & \hat{f}_i(R_{1,2}) & \cdots & \hat{f}_i(R_{1,x_R}) \\ \vdots & & \ddots & \vdots \\ \hat{f}_i(R_{x_R,1}), & \hat{f}_i(R_{x_R,2}) & \cdots & \hat{f}_i(R_{x_R,x_R}) \end{bmatrix} \qquad \text{(Eq. 3-12)}$$

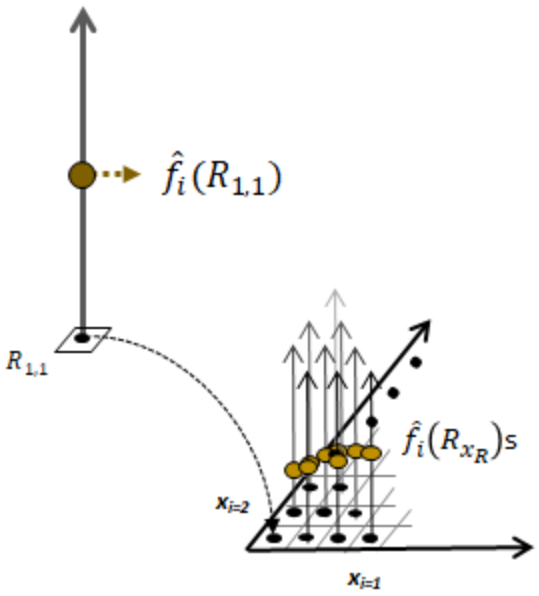


Figure 3 - 4, A SOI Operational Environment with Mean-Response Point Estimates

3.2.4.3 Generate Element Performance Prediction Confidence Intervals

3.2.4.3.1 Calculate Elements' Point Estimate Variances

With the employment of least-square estimation method, each of $\widehat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$'s calculated matrix of coefficients, $\widehat{\boldsymbol{b}}_{\boldsymbol{f_i}}$, is considered to be normally distributed with mean vector $\boldsymbol{\beta}_{\boldsymbol{f_i}}$ and covariance matrix of $\boldsymbol{Var}[\widehat{\boldsymbol{f}_{\boldsymbol{i}}(\boldsymbol{R}_{\boldsymbol{x_R}})}] = {\sigma_{\boldsymbol{f_i}}}^2 \cdot (\boldsymbol{X}_{\boldsymbol{f_i}}{}'\boldsymbol{X}_{\boldsymbol{f_i}})^{-1}$ (Myers & Montgomery, 2002). The unknown ${\sigma_{\boldsymbol{f_i}}}^2$ is typically substituted with the captured raw performance data's ANOVA-generated Mean Square Error (MS_E) as defined in Eq. 3-13 (Myers & Montgomery, 2002, p. 41).

$$\hat{\sigma}_{f_i}{}^2 = MS_E = {SS_E}/{df_E} = (\boldsymbol{y}_{f_i} - \mathbf{X}_{f_i} \cdot \widehat{\boldsymbol{b}}_{f_i})' \cdot (\boldsymbol{y}_{f_i} - \mathbf{X}_{f_i} \cdot \widehat{\boldsymbol{b}}_{f_i}) \Big/ N - (k+1) \quad \text{(Eq. 3-13)}$$

Based on this defined normal-distribution relationship of $\boldsymbol{N(\beta_{f_i}, \widehat{Var[f_i(R_{x_R})]})}$, each individual regression coefficient within $\hat{\boldsymbol{b}}_{\boldsymbol{f_i}}$, $\hat{b}_j$s, can be further characterized with a corresponding student *t*-based $100 \cdot (1 - \textit{predetermined confidence coefficient, } \alpha)\%$ confidence interval as statistically expressed in Eq. 3-14.

$$\hat{b}_j - t_{\frac{\alpha}{2},n-p}\sqrt{\hat{\sigma}_{\boldsymbol{f}_i}{}^2 C_{jj}} \leq \hat{b}_j \leq \hat{b}_j + t_{\frac{\alpha}{2},n-p}\sqrt{\hat{\sigma}_{\boldsymbol{f}_i}{}^2 C_{jj}}, \quad j = 0,1,\dots,k,$$

where Cjj is the (j, j)th component of the matrix $(\boldsymbol{X_{f_i}}' \cdot \boldsymbol{X_{f_i}})^{-1}$, *and*

$$\hat{b}_j - t_{\frac{\alpha}{2},n-p} se(\hat{b}_j) \leq \hat{b}_j \leq \hat{b}_j + t_{\frac{\alpha}{2},n-p} se(\hat{b}_j), \quad j = 0,1,\dots,k$$

where the standard error of b_j, $se(b_j) = \sqrt{\sigma^2 \cdot C_{jj}}$ (Eq. 3-14)

The variations in each regression coefficient of $\hat{b}_j$, as contributed from variations of the corresponding *Cjj* component within the matrix $(\boldsymbol{X_{f_i}}' \boldsymbol{X_{f_i}})^{-1}$, collectively generate the cumulative variations in the element mean-response model, $\hat{\boldsymbol{f}}_i(\mathbf{R})$. Therefore, a corresponding $100 \cdot (1 - \textit{predetermined confidence coefficient, } \alpha)\%$ prediction confidence-interval at every R_{x_R} subspace, $\hat{f}_{i_{variability}}(R_{x_R}) = \{\hat{f}_i(R_{x_R})_{min}, \hat{f}_i(R_{x_R})_{min}\}|\alpha$, can be defined as shown in Eq. 3-15 (Montgomery, 2009, p. 407; Myers & Montgomery, 2002, p 41). This *t*-based confidence interval for the mean, $\hat{f}_i(R_{x_R})$, is further illustrated in Figure 3-5.

$$\hat{f}_i(\mathbf{x_0}) - \mathrm{t}_{\frac{\alpha}{2},\mathrm{n-p}} \cdot \sqrt{\hat{\sigma}_{f_i}{}^2 \cdot \sqrt{Var[\widehat{f_i(R_{x_R})}]}} \leq \hat{f}_i(\mathbf{x_0})$$

$$\leq \hat{f}_i(\mathbf{x_0}) + \mathrm{t}_{\frac{\alpha}{2},\mathrm{n-p}} \cdot \sqrt{\hat{\sigma}_{f_i}{}^2 \cdot \sqrt{Var[\widehat{f_i(R_{x_R})}]}}$$

where $R_{x_R} \overset{\text{def}}{=} \boldsymbol{x}_0$*, and* $\hat{\sigma}_{f_i}$ *is defined as* $MS_E = \left. \sum_i^N (y_i - \bar{y}_n) \right/ df_{MSE}$ (Myers & Montgomery, 2002, 318), *and*

$$\hat{f}_i(\mathbf{x_0}) - \mathrm{t}_{\frac{\alpha}{2},\mathrm{n-p}} \cdot \sqrt{\hat{\sigma}_{f_i}{}^2 \cdot \left(1 + \mathbf{x}_0{}' \left(\boldsymbol{X}_{\boldsymbol{f_i}}{}' \boldsymbol{X}_{\boldsymbol{f_i}}\right)^{-1} \mathbf{x}_o\right)} \leq \hat{f}_i(\mathbf{x_0})$$

$$\leq \hat{f}_i(\mathbf{x_0}) + \mathrm{t}_{\frac{\alpha}{2},\mathrm{n-p}} \cdot \sqrt{\hat{\sigma}_{f_i}{}^2 \cdot (1 + \mathbf{x}_0{}' \left(\boldsymbol{X}_{\boldsymbol{f_i}}{}' \boldsymbol{X}_{\boldsymbol{f_i}}\right)^{-1} \mathbf{x}_0)}, \textit{ or}$$

$$\hat{f}_i(R_{x_R})_{min} \leq \hat{f}_{i_variability}(R_{x_R}) \leq \hat{f}_i(R_{x_R})_{min} \qquad \text{(Eq. 3-15)}$$

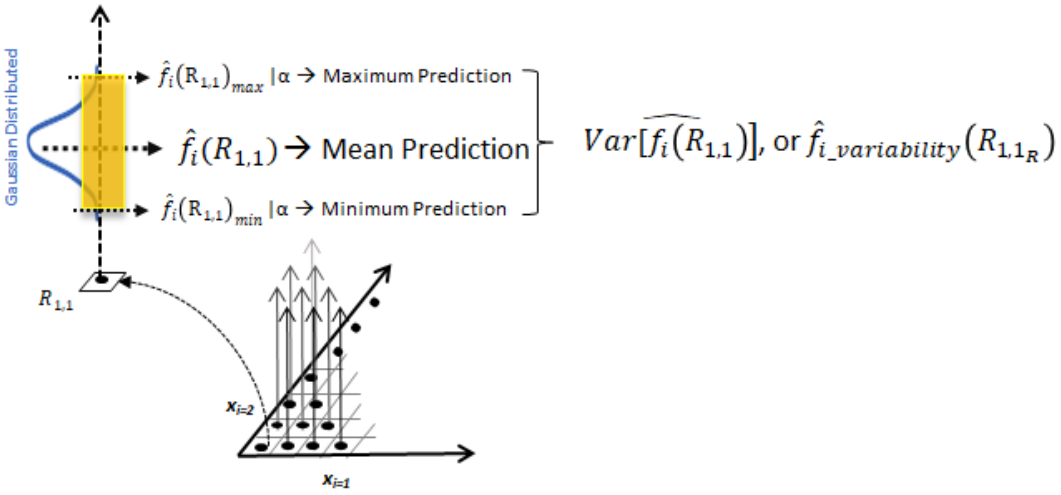


Figure 3 - 5, Calculated Prediction Confidence-Interval Point Estimate

3.2.4.3.2 Construct Elements' Prediction-Variability Models

Once element prediction confidence-interval point estimates, $\hat{f}_{i_variability}(R_{x_R})$s, are individually derived for all $\mathrm{R}_{\mathrm{x_R}}$ subspaces, they are to be consolidated into the *k*-

dimensional numerical array of $\boldsymbol{R}$ to construct the respective element prediction-confidence interval response model, $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$. Mathematical expressions of constructed $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$ within a (k=2)-dimensional numerical array of $\boldsymbol{R}$ from Eq. 3-4 is shown in Eq. 3-16 and illustrated in Figure 3-6 as a collective construct of Figure 3-5.

$$\hat{\boldsymbol{f}}_{i_{variability}}(\boldsymbol{R})\,|\alpha = \{\hat{\boldsymbol{f}}_i(\boldsymbol{R})_{min}, \hat{\boldsymbol{f}}_i(\boldsymbol{R})_{minmax}\}|\alpha$$

$$= \begin{bmatrix} \hat{f}_i(R_{1,1})_{min}, \hat{f}_i(R_{1,1})_{max}\}|\alpha & \cdots & \hat{f}_i(R_{1,x_R})_{min}, \hat{f}_i(R_{1,x_R})_{max}\}|\alpha \\ \vdots & \ddots & \vdots \\ \hat{f}_i(R_{x_R,1})_{min}, \hat{f}_i(R_{x_R,1})_{max}\}|\alpha & \cdots & \hat{f}_i(R_{x_R,x_R})_{min}, \hat{f}_i(R_{x_R,x_R})_{max}\}|\alpha \end{bmatrix}$$

where $\hat{f}_i(R_{r,c})_{min}, \hat{f}_i(R_{r,c})_{max}\}|\alpha = \hat{f}_i(R_{r,c}) \pm t_{\frac{\alpha}{2},n-p} \cdot \sqrt{\hat{\sigma}_{f_i}{}^2(1 + \mathbf{x}_{R_{r,c}}{}' (\mathbf{X}'\mathbf{X})^{-1} \cdot \mathbf{x}_{R_{r,c}})}$

where $\boldsymbol{X} \equiv \boldsymbol{X}_{\boldsymbol{f_i}}$ *and* $\boldsymbol{x}_{R_{l,m}}$ *depicts the moment matrix corresponding to the* $R_{l,m}$ *subspace.*

(Eq. 3-16)

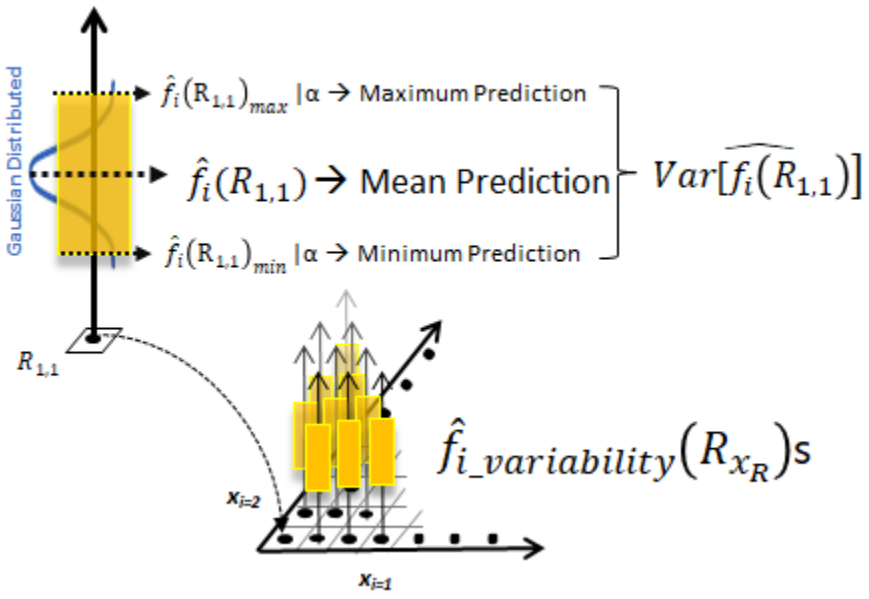


Figure 3 - 6, SOI Operational Environment with Confidence-Interval Point Estimates

At this juncture of SBED, every element-under-test's captured and collected element performance and verification data (from Step 3) are now modeled as respective $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$s and $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$s that are collectively normalized w.r.t. the SOI operational environment, $\boldsymbol{R}(\xi_j\text{s})$ (as defined in Eq. 3-2). Thus, RBT&E test approach's existing data-domain inconsistency challenge has been effectively resolved. However, and as pertain to the data-metrological inconsistency challenge, the output response variables of each elements' respective $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$ and $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$ are still individually presented as mismatched measurement units or response outputs specific to their allocated requirements' variables (e.g., element A's response models are outputted as a function in *mph* speed, element B's as a function in y_2-tons of payload, element C's as a function in y_3-*Watts* of sustained output power, etc.). All the calculated $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$s and $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$s still need to be collectively standardized to a common output variable to support their subsequent quantitative amalgamation to produce an objective SOI-capability forecast.

3.2.4.4 Infer Element Capability Models

SBED's aim is to identify and introduce a common SOI mission-representative data unit-of-measurement to standardize all (previously mismatched) RBT&E element verification data. Knowing that decision makers and the T&E community both desire to assess/evaluate element performances w.r.t. their allocated requirements' thresholds (i.e., the *Requirement-Based Performance Verification T&E* goal per INCOSE [2015] and DoD [2017]), SBED adopts the *"predicted performance-versus-requirement threshold"* assessment criteria as the common unit for all response output standardization. When elements' $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$s and $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$s are remodeled as the standardized *"predicted*

performance-versus-requirement threshold" output response variable that are influenced by the normalized SOI operational environment, they subsequently become the respective (and now matching) *element-capability* forecasts, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capablity}}(\boldsymbol{R})$s.

3.2.4.4.1 Calculate Element-Capability Point Estimates

Each of SBED's *element-capability* forecast, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capablity}}(\boldsymbol{R})$, is also a *k*-dimensional construct of $(x_R)^k$ number of respective scalar *element-capability* point estimates, $\hat{f}_{i_capability}(R_{x_R})$s. One at every R_{x_R} subspace throughout the entire $\boldsymbol{R}$. Each $\hat{f}_{i_capability}(R_{x_R})$ is individually calculated through the quantitative comparison between: (a) the element's respective $\hat{f}_{i_variability}(R_{x_R})$'s boundaries and/or $\hat{f}_i(R_{x_R})$, and (b) the allocated element requirements' stated threshold values, T_{f_i}, or $T_{f_{i_R_{x_R}}}$,if the requirements include dynamic threshold values.

SBED proposes two quantitative comparison methods to obtain $\hat{f}_{i_capability}(R_{x_R})$ and they are explained in the following paragraphs. As long as all elements within the set $\boldsymbol{G}$ (previously defined in SBED Step 1's Eq. 3-1) systematically derive their corresponding $\hat{f}_{i_capability}(R_{x_R})$s through one of these two comparison methods, domain-metrological matching *element-capability* forecasts, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capablity}}(\boldsymbol{R})$s, can be obtained.

The first comparison method to obtain $\hat{f}_{i_capability}(R_{x_R})$ is *probability of meeting requirement threshold* (hereon referred to as the "probability format") as explained in Eq. 3-17. It subjects each scalar prediction confidence interval's min and max boundaries, $\hat{f}_i(R_{x_R})_{min}$ and $\hat{f}_i(R_{x_R})_{max}$ from Eq. 3-15, to either a one- or two-tail statistical test. Treating the respective mean-response point estimate, $\hat{f}_i(R_{x_R})$, as μ and the corresponding

element requirement's T_{f_i} or $T_{f_i_R_{x_R}}$ as the null value. The comparison method then interprets the complement of area under the rejection region, $1 - \hat{p}\left(f_i \; violating \; T_{R_{x_R}} \middle| R_{x_R}\right)$, as R_{x_R} subspace's respective *element-capability* point estimate, $\hat{f}_{i_capability}(R_{x_R})$. This statistical comparison between elements' respective $\hat{f}_{i_variability}(R_{x_R})$'s boundaries and/or $\hat{f}_i(R_{x_R})$ and allocated requirement thresholds is further illustrated in Figure 3-7.

if: $\hat{f}_i(R_{x_R})_{min} < T_{R_{x_R}}$*, then:*

$\hat{p}\left(f_i \; violating \; T_{R_{x_R}} \middle| R_{x_R}\right) =$

1-normcdf$\{(\hat{f}_i(R_{x_R}) - T_i)/(\hat{\sigma}_{\boldsymbol{f_i}} \cdot \sqrt{(\boldsymbol{x}_{R_{x_R}} \cdot (\boldsymbol{X_{f_i}}' \cdot \boldsymbol{X_{f_i}})^{-1} \cdot \boldsymbol{x}_{R_{x_R}})})\}$*, and*

$\hat{f}_{i_capability}(R_{x_R}) = 1 - \hat{p}\left(f_i \; violating \; T_{R_{x_R}} \middle| R_{x_R}\right)$*;*

else if: $\hat{f}_i(R_{x_R})_{max} > T_{R_{x_R}}$*, then:*

$\hat{p}\left(f_i \; violating \; T_{R_{x_R}} \middle| R_{x_R}\right) =$

normcdf$\{(\hat{f}_i(R_{x_R}) - T_i)/(\hat{\sigma}_{\boldsymbol{f_i}} \cdot \sqrt{(\boldsymbol{x}_{R_{x_R}} \cdot (\boldsymbol{X_{f_i}}' \cdot \boldsymbol{X_{f_i}})^{-1} \cdot \boldsymbol{x}_{R_{x_R}})})\}$*, and*

$\hat{f}_{i_capability}(R_{x_R}) = 1 - \hat{p}\left(f_i \; violating \; T_{R_{x_R}} \middle| R_{x_R}\right)$*;*

else (no failure predicted), then:

$\hat{p}\left(f_i \; violating \; T_{R_{x_R}} \middle| R_{x_R}\right)$ *=0%, and*

$\hat{f}_{i_capability}(R_{x_R}) = 1 - \hat{p}\left(f_i \; violating \; T_{R_{x_R}} \middle| R_{x_R}\right) = 100\%$

where 'normcdf" is the normal cumulative distribution function. (Eq. 3-17)

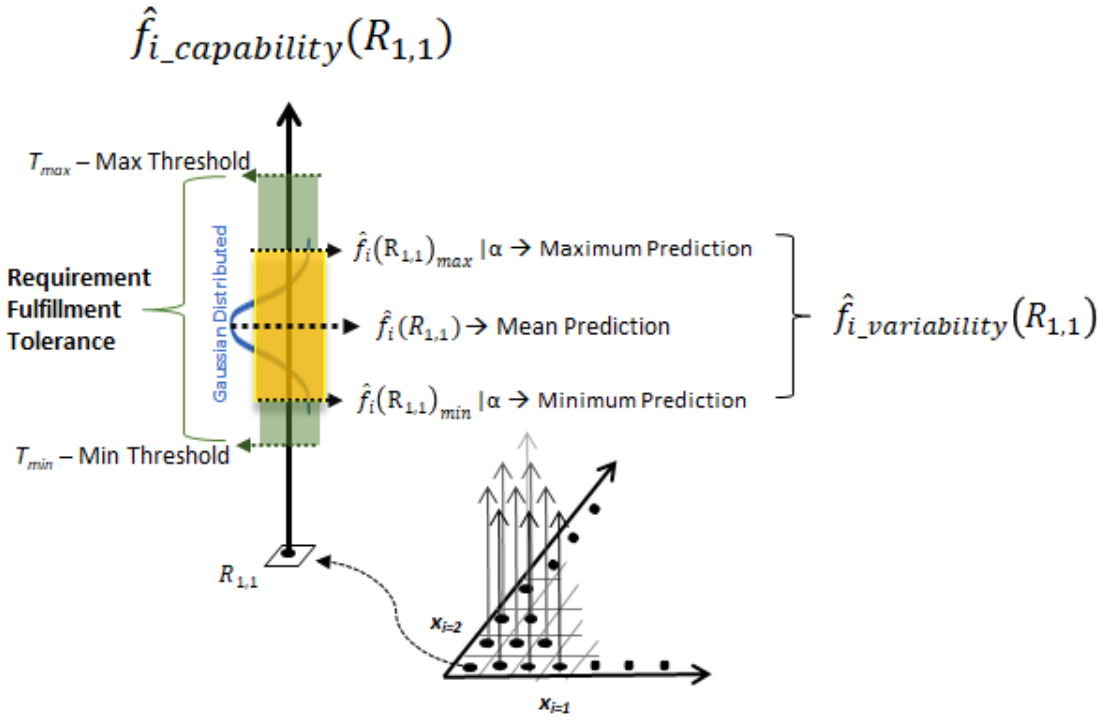


Figure 3 - 7, Derived Element Capability Point Estimate via *t*-Test

The second quantitative comparison method is a discrete *meeting or violating requirement threshold* indicator (hereon referred to as the "discrete format"), explained in Eq. 3-18. It directly compares each scalar mean-response point estimate, $\hat{f}_i(R_{x_R})$, against the corresponding T_{f_i} or $T_{f_{i_}R_{x_R}}$, and then interprets the discrete comparison result as the R_{x_R} subspace's respective *element-capability* point estimate in discrete format.

$$\textit{if: } (\hat{f}_i(R_{x_R}) < T_{R_{x_R},min}) \text{ or } (\hat{f}_i(R_{x_R}) > T_{R_{x_R},max}),$$

$$\textit{then: } \hat{f}_{i_capability}(R_{x_R}) \overset{\text{def}}{=} \boldsymbol{meeting}\ T_{R_{x_R}}|R_{x_R}$$

$$\textit{else: } \hat{f}_{i_capability}(R_{x_R}) \overset{\text{def}}{=} \boldsymbol{violating}\ T_{R_{x_R}}|R_{x_R}. \qquad \text{(Eq. 3-18)}$$

3.2.4.4.2 Construct Element-Capability Models

Once all $(x_R)^k$ number of *element-capability* point estimates, $\hat{f}_{i_capability}(R_{x_R})$s, are calculated in either probability or discrete format, they are to be consolidated into the *k*-dimensional numerical array of $\boldsymbol{R}$ to construct the respective *element-capability* forecast, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$. Mathematical expressions of the constructed $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$ within a (*k*=2)-dimensional $\boldsymbol{R}$ as defined in Eq. 3-4 are shown in Eq. 3-19 and Eq. 3-20 in probability and discrete formats respectively. The constructed $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$ within a (*k*=2)-dimensional $\boldsymbol{R}$, in probability format, is also illustrated in Figure 3-8 as a collective construct of Figure 3-7.

$$\hat{\boldsymbol{f}}_{\boldsymbol{i_{capability}}}(\boldsymbol{R})$$

$$= \begin{bmatrix} 1-\hat{\mathrm{p}}\left(f_i\ \boldsymbol{violating}\ \mathrm{T}_{\mathrm{R}_{1,1}} \middle| \mathrm{R}_{1,1}\right) & \cdots & 1-\hat{\mathrm{p}}\left(f_i\ \boldsymbol{violating}\ \mathrm{T}_{\mathrm{R}_{1,x_R}} \middle| \mathrm{R}_{1,x_R}\right) \\ \vdots & \ddots & \vdots \\ 1-\hat{\mathrm{p}}\left(f_i\ \boldsymbol{violating}\ \mathrm{T}_{\mathrm{R}_{x_R,1}} \middle| \mathrm{R}_{x_R,1}\right) & \cdots & 1-\hat{\mathrm{p}}\left(f_i\ \boldsymbol{violating}\ \mathrm{T}_{\mathrm{R}_{x_R,x_R}} \middle| \mathrm{R}_{x_R,x_R}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \hat{\mathrm{p}}\left(f_i\ \boldsymbol{meeting}\ \mathrm{T}_{\mathrm{R}_{1,1}} \middle| \mathrm{R}_{1,1}\right) & \cdots & \hat{\mathrm{p}}\left(f_i\ \boldsymbol{meeting}\ \mathrm{T}_{\mathrm{R}_{1,x_R}} \middle| \mathrm{R}_{1,x_R}\right) \\ \vdots & \ddots & \vdots \\ \hat{\mathrm{p}}\left(f_i\ \boldsymbol{meeting}\ \mathrm{T}_{\mathrm{R}_{x_R,1}} \middle| \mathrm{R}_{x_R,1}\right) & \cdots & \hat{\mathrm{p}}\left(f_i\ \boldsymbol{meeting}\ \mathrm{T}_{\mathrm{R}_{x_R,x_R}} \middle| \mathrm{R}_{x_R,x_R}\right) \end{bmatrix} \quad \text{(Eq. 3-19)}$$

$$\hat{\boldsymbol{f}}_{\boldsymbol{i_capabiltiy}}(\boldsymbol{R}) = \begin{bmatrix} \hat{f}_{i_{capability}}{}^{*} \left|\mathrm{R}_{1,1}\right. & \cdots & \hat{f}_{i_{capability}}{}^{*} \left|\ \mathrm{R}_{1,x_R}\right. \\ \vdots & \ddots & \vdots \\ \hat{f}_{i_{capability}}{}^{*} \left|\mathrm{R}_{x_R,1}\right. & \cdots & \hat{f}_{i_{capability}}{}^{*} \left|\ \mathrm{R}_{x_R,x_R}\right. \end{bmatrix} \quad \text{(Eq. 3-20)}$$

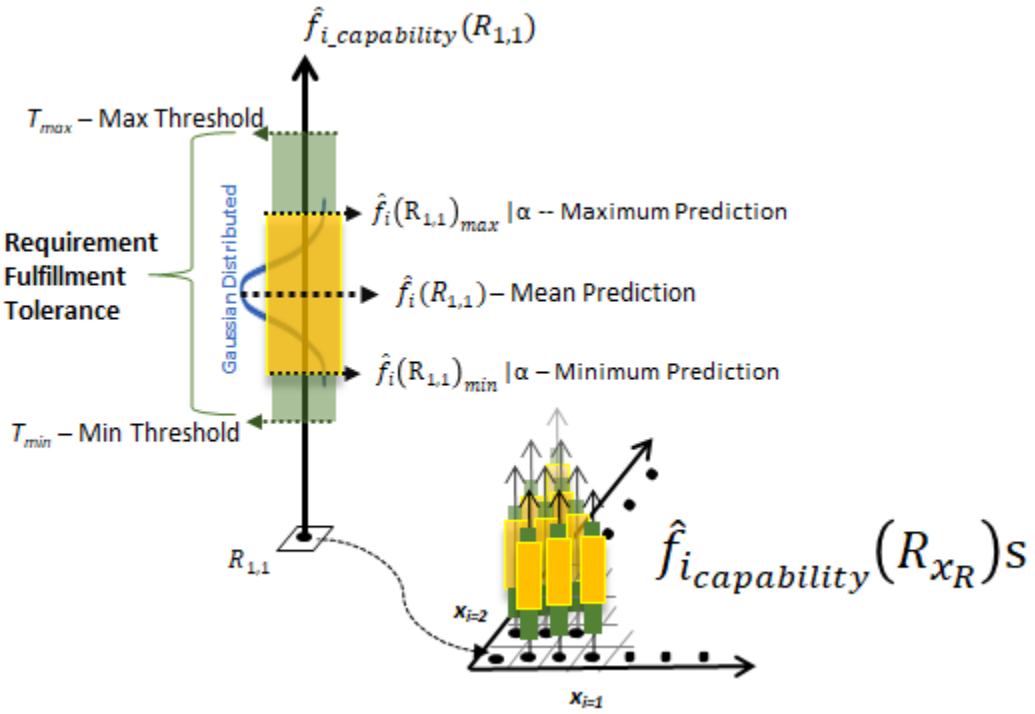


Figure 3 - 8, Calucated *Element-Capability* Point Estimate

As a visual aid of SBED's $\hat{f}_{i_capability}(R_{x_R})$ quantitative-comparison methodology in probability format, illustrations of a hypothetical $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$ and $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$ within a *(k=2)*-dimensional $\boldsymbol{R}$ are shown in Figure 3-9. Figure 3-9 also includes illustrations of the allocated requirement's (no greater than) threshold, T_{f_i}. Their respective *element-capability* forecast, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$ in probability format, is shown in Figure 3-10 as a contour plot consisting of all R_{x_R} subspaces' *element-capability* point estimates, $\hat{f}_{i_capability}(R_{x_R})$s, or the statistical-comparison results between R_{x_R} subspaces' $\hat{f}_i(R_{x_R})_{max}$s and the requirement's $\leq T_{f_i}$, $1-\hat{p}\left(f_i\ violating\ T_{R_{x_R}} \middle| R_{x_R}\right)$s per Eq. 3-17.

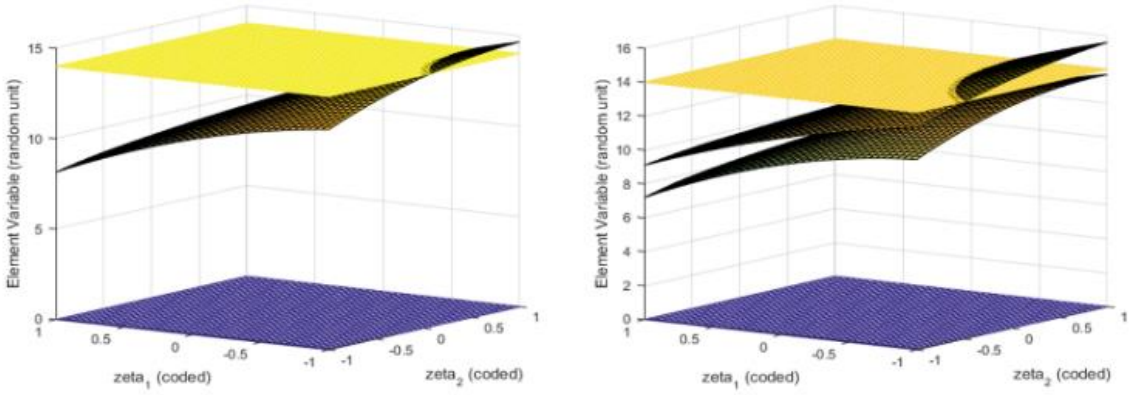


Figure 3 - 9, Examples of SBED's Element Response Plots

Figure 3-9. (a) and (b) MATLAB surface plots of element's SBED-produced $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$ and $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$ respectively shown with the element's $\boldsymbol{T}_{\boldsymbol{f_i}}$ (yellow).

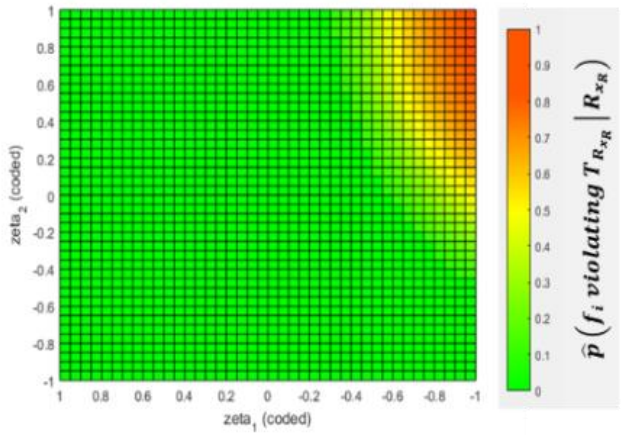


Figure 3 - 10, Contour plot SBED-Producred Element-Capablity Responses

Note. Derived from comparing its $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R}_{\boldsymbol{x_R}})_{\boldsymbol{max}}$s against (< $\boldsymbol{T}_{\boldsymbol{f_i}}$) in probability format.

Once again, and as previously discussed in Subsection 3.2.3.3 (Alignment to the Primary Element Verification's Goals and Objectives), it should be noted that every element's respective regression models, $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$, $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$, and $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$, are fitted prediction model that may include additional parameters other than those that define

R (e.g., parameters unique only to the allocated element requirement). During the subsequent SOI-capability estimation that ensues in SBED Step 5, these non-***R*** parameters values shall be kept constant at their element-requirement specified values (i.e., considered to be inert).

3.2.4.5 Resolution of RBT&E's Data-Inconsistency Challenges

At the completion of SBED's Step 4, note that through its delivery of collectively standardized and normalized *element-capability* regression models, SBED has successfully introduced a resolution to the SOI Implementation process' existing RBT&E data-metrological and –domain inconsistency challenges. This dissertation uses the arbitrary to-be-forecasted SOI-capability that is originally introduced in Subsection 3.2.1 (Depict the SOI-Capability Scenario), Figure 3-1, as a brief hypothetical demonstration.

First, SBED determines that the characterization of each SOI capability's four element verification results are to be normalized w.r.t. two arbitrarily selected SOI operational environmental parameters designated as ξ_A, ξ_B (as coded x_A and x_B between {-1, +1} and unitless per Eq. 3-3) as illustrated in Figure 3-1). Then through the employment of various DoE designs, SBED captures the effects on the four elements' respective performances w.r.t. changes throughout the entire SOI operational environment. (Note that in this hypothetical demonstration's effort for realism, multiple raw experimental laboratory data sets from various independent sources are used. They are Richert, Morr, & Cooney [1974], Myers & Montgomery [2002], and El-Say [2016]. Descriptions of each employed DoE design are listed in Table 3-3 below and the original raw data sets from the original independent sources are presented in Appendix A.) Subsequently, each of the four

element verification result's respective linear regression mean-response models, $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$s, *for i = 1 to 4,* and prediction confidence-interval models, $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$, *for i =1 to 4,* are computed as 2nd-order polynomials (via Minitab data-analysis software) based on: the actual organic behavior of the independent sources' experimental entity-under-investigation per Richert et al. (1974), Myers & Montgomery (2002), and El-Say (2016), and the extensiveness of their DoE design(s) employed to capture that behavior. The mean response models are presented in Eq. 3-21 below, and both mean-response and prediction confidence-interval models are illustrated from Figure 3-11 to Figure 3-14.

Table 3 - 3, DoE Methods' Reported Verification Data

Element Requirement #	Employed DoE Designs	Sample Size	Controlled Factors [A]
$\boldsymbol{f}_1$	2^{5-1} Factorial + Central Composite Design [B]	$16_{screening}$ + $14_{sequential}$	$\boldsymbol{x_A, x_B, x_1, x_2, x_3}$ (coded & unitless)
$\boldsymbol{f}_2$			
$\boldsymbol{f}_3$	Box Behnken Design [C]	15	$\boldsymbol{x_A, x_B, x_4}$ (coded & unitless)
$\boldsymbol{f}_4$	Small Composite Design (saturated) [D]	18	$\boldsymbol{x_A, x_B, x_5, x_6,}$ (coded & unitless)

Note A: factors with numeric subscripts indicate parameters intrinsic only to the respective element requirement's verification condition.
Note B: DoE data from Reichert, et al. (1974).
Note C: DoE data from Myers & Montgomery (2002).
Note D: DoE data from El-Say (2016).

$$\hat{\boldsymbol{f}}_{\boldsymbol{1}}(\boldsymbol{R}) = 11.851 - 1.761 \cdot \zeta_1 + 0.576 \cdot \zeta_2 - 0.182 \cdot x_1 + 0.4430 \cdot x_2 + 0.191 \cdot x_3 - 0.634 \cdot \zeta_1{}^2 - 0.512 \cdot \zeta_2{}^2 - 0.302 \cdot x_1{}^2 - 0.067 \cdot x_2{}^2 - 0.156 \cdot x_3{}^2 - 0.489 \cdot \zeta_1 \cdot \zeta_2 - 0.250 \cdot \zeta_1 \cdot x_1 - 0.122 \cdot \zeta_1 \cdot x_2 - 0.368 \cdot \zeta_1 \cdot x_3 - 0.029 \cdot \zeta_2 \cdot x_1 - 0.044 \cdot \zeta_2 \cdot x_3 + 0.24 \cdot \zeta_2 \cdot x_3 + 0.167 \cdot x_1 \cdot x_2 + 0.11 \cdot x_1 \cdot x_3 - 0.01 \cdot x_2 \cdot x_3$$

$$\hat{\boldsymbol{f}}_{\boldsymbol{2}}(\boldsymbol{R}) = 77.79 - 10.121 \cdot \zeta_1 - 0.104 \cdot \zeta_2 - 8.679 \cdot x_1 - 0.712 \cdot x_2 + 1.371 \cdot x_3 - 4.157 \cdot {\zeta_1}^2 - 2.657 \cdot \zeta^2 - 4.769 \cdot {x_1}^2 - 0.344 \cdot {x_2}^2 - 0.332 \cdot {x_3}^2 + 2.77 \cdot \zeta_1 \cdot \zeta_2 + 6.21 \cdot \zeta_1 \cdot x_1 + 1.77 \cdot \zeta \cdot x_2 + 0.71 \cdot \zeta_1 \cdot x_3 - 1.68 \cdot \zeta_2 \cdot x_1 + 0.04 \cdot \zeta_2 \cdot x_2 + 1.51 \cdot \zeta_2 \cdot x_3 - 0.26 \cdot x_1 \cdot x_2 + 0.48 \cdot x_1 \cdot x_3 + 0.03 \cdot x_2 \cdot x_3$$

$$\hat{\boldsymbol{f}}_{\boldsymbol{3}}(\boldsymbol{R}) = 62.00 + 1.00 \cdot \zeta_1 + 2.63 \cdot \zeta_2 - 2.38 \cdot x_4 - 7.37 \cdot {\zeta_1}^2 + 1.88 \cdot {\zeta_2}^2 - 3.63 \cdot {x_4}^2 - 2.00 \cdot \zeta_1 \cdot \zeta_2 + 11.00 \cdot \zeta_1 \cdot x_4 + 1.75 \cdot \zeta_2 \cdot x_4$$

$$\hat{\boldsymbol{f}}_{\boldsymbol{4}}(\boldsymbol{R}) = 25.51 - 1.49 \cdot \zeta_1 + 0.06 \cdot \zeta + 6.471 \cdot x_5 + 2.23 \cdot x_6 - 1.08 \cdot {\zeta_1}^2 - 0.55 \cdot \zeta^2 + 0.21 \cdot {x_5}^2 - 0.53 \cdot {x_6}^2 - 1.28 \cdot \zeta_1 \cdot \zeta_2 - 1.63 \cdot \zeta_1 \cdot x_5 - 1.72 \cdot \zeta_1 \cdot x_6 - 2.36 \cdot \zeta_2 \cdot x_5 - 3.95 \cdot \zeta_2 \cdot x_6 - 1.98 \cdot x_5 \cdot x_6$$

(Eq. 3-21)

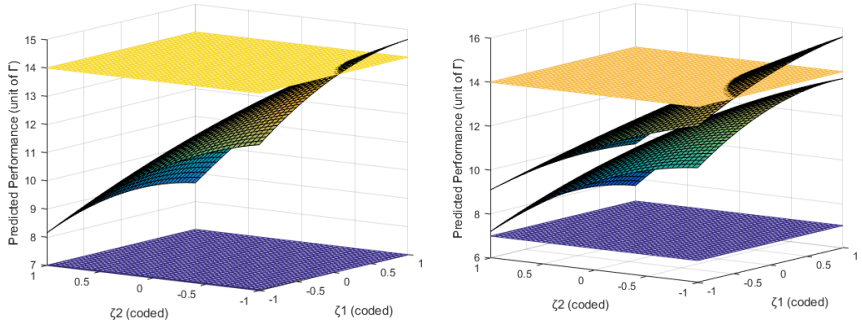


Figure 3 - 11, Element f_1's Response Models

Figure 3-11. Element f_1's mean-response model (left) and prediction confidence-interval model (right) with the allocated requirement's thresholds, *Tmax* and *Tmin* (from Table 2-1) shown as yellow and purple planes respectively. Both models are as influenced by the SOI operational environment, $R(\zeta_A, \zeta_B)$, with remaining parameters, x_1B, and $x_2\Gamma$, held constant w.r.t. the requirement's specified verification condition (from Table 2-1).

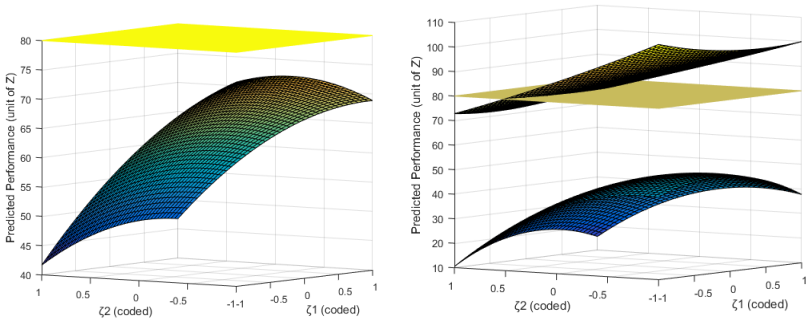


Figure 3 - 12, Element f_2's Response Models

Figure 3-12. Element f_2's mean-response model (left) and prediction confidence-interval model (right), with the allocated requirement's threshold T_2 (from Table 2-1) shown in yellow. Both models are as influenced by the SOI operational environment, $R(\zeta_A, \zeta_B)$, with remaining parameters, x_1B and x_3E, held constant w.r.t. the requirement's specified verification condition (from Table 2-1).

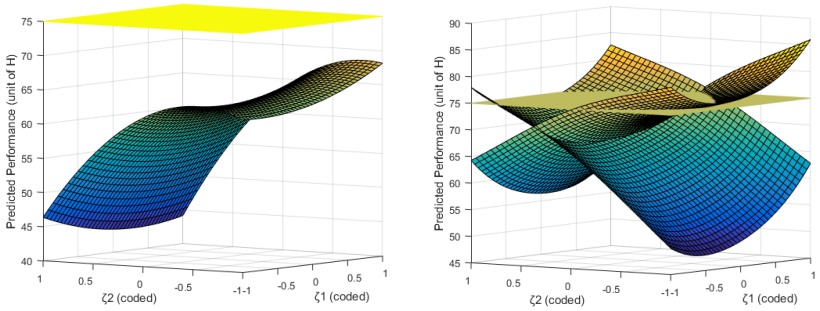
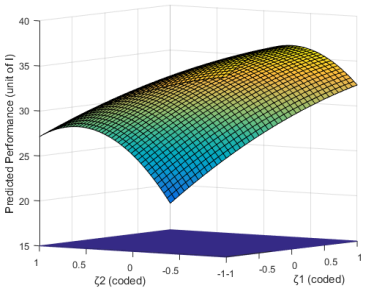


Figure 3 - 13, Element f_3's Response Models

Figure 3-13. Element f_3's mean-response model (left) and prediction confidence-interval model (right), with the allocated requirement's threshold T_3 (from Table 2-1) shown in yellow. Both models are as influenced by the SOI operational environment, $R(\zeta_A, \zeta_B)$, with remaining parameter x_4Z held constant w.r.t. the requirement's specified verification condition (from Table 2-1).

Note. The intersecting maximum and minimum prediction confidence-interval boundaries are due to the swap between calculations using $x_4(min)$ and $x_4(max)$.



Not useable due to the incomplete (saturated) DoE design employed.

Figure 3 - 14, Element f_4's Response Models

Figure 3-14. Element f_4's mean-response model with the allocated requirement's threshold T_4 (from Table 2-1) shown in purple. The model is as influenced by the SOI operational environment, $R(\zeta_A, \zeta_B)$, with remaining parameters, $x_5\Theta$ and $x_6 I$, held constant w.r.t. the requirement's specified verification condition (from Table 2-1).
Note. Element f_4's prediction confidence-interval model is not presented due to the employment of a saturated DoE design (El-Say, 2016).

Finally, each of the four element's (now normalized and standardized) *capability* response models are computed through the statistical significance tests using the allocate element requirement thresholds, T_is, as the null value. Note that these threshold values are hypothetical as they are extracted from Subsection 2.2.4 and Tables 2-1's descriptions. Through SBED's element-capability probability-comparison technique, per Eq. 3-17, the respective element prediction confidence-interval models' appropriate boundaries (i.e., maximum or minimum) are subjected to either one-tail (e.g., element $\boldsymbol{f_1}$) or two-tail statistical tests (e.g. elements $\boldsymbol{f_2}$, $\boldsymbol{f_3}$, and $\boldsymbol{f_4}$) while treating the respective element mean-response models (from Eq. 3-21) as μ_{f_i}s. The corresponding ($k=2$)-dimensional mesh plots for these four *element-capability* response models (similar to Figure 3-8 and generated via MATLAB) are shown in Figure 3-15. The scope of each SBED-produced normalized and standardized *element capability* models, or the ability to achieve element requirements

throughout the entire SOI operational environment, $\boldsymbol{R} \stackrel{\text{def}}{=} \{\xi_A, \xi_B\}$ or $\boldsymbol{R} \stackrel{\text{def}}{=} \{x_A, x_B\}$) as defined in Table 3-4.

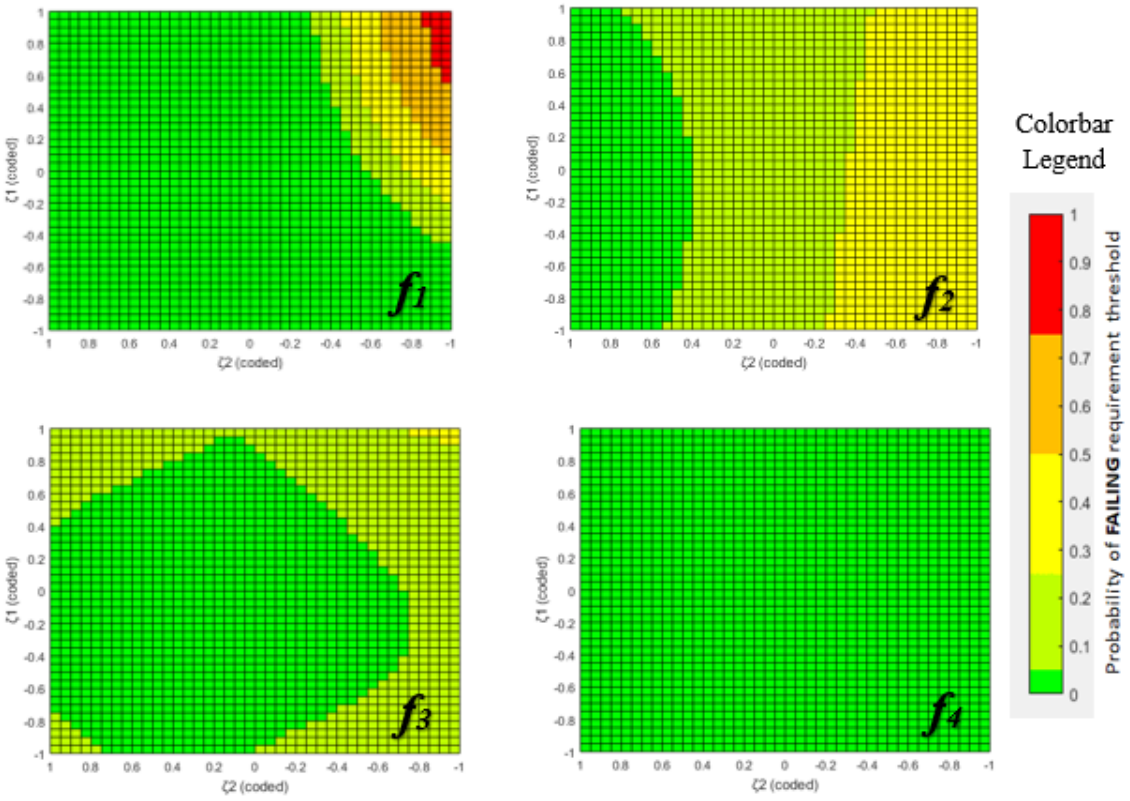


Figure 3 - 15, Example's four element capability models.

Figure 3-15, all response models are as influenced by the system's operational environment, R(ζ1, ζ2), with remaining parameters held constant w.r.t. the requirement's specified verification condition (from Table 2-2).

Table 3 - 4, Explanations of *Element-Capability* Response Models

Element Requirement #	***Element Capability* Output Definition**
$\boldsymbol{f}_1$	$P(7.5 \leq \boldsymbol{f_1} \leq 15\ \boldsymbol{A}) \mid x_A, x_B, \boldsymbol{x_1}^*, \boldsymbol{x_2}^*, \boldsymbol{x_3}^*$
$\boldsymbol{f}_2$	$P(\boldsymbol{f_2} \leq \boldsymbol{80\ \Delta}) \mid x_A, x_B, \boldsymbol{x_1}^*, \boldsymbol{x_2}^*, \boldsymbol{x_3}^*$
$\boldsymbol{f}_3$	$P(\boldsymbol{f_3} \leq \boldsymbol{75\ E}) \mid x_A, x_B, \boldsymbol{x_4}^*$
$\boldsymbol{f}_4$	$P(\boldsymbol{f_4} \geq \boldsymbol{15\ H}) \mid x_A, x_{B,}, \boldsymbol{x_5}^*, \boldsymbol{x_6}^*,$

Note: * Represents scalar conditional variables.

By directly comparing any one of the RBT&E's reporting formats in Table 3-5 against SBED's objectively-produced element mean response models, $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(\boldsymbol{R})$s, element mean response prediction models, $\hat{\boldsymbol{f}}_{\boldsymbol{i_variability}}(\boldsymbol{R})$s, and/or *element-capability* forecast models, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$s (e.g., Figure 3-11 to Figure 3-14), it is self-evident that SBED provides far more descriptive and comprehensive visualizations of captured element performances. Instead of the traditional RBT&E approach's presentation of demonstrated performance and deficiencies w.r.t. only the requirement verification conditions (i.e., Table 3-5 'Test-Condition' Column), the SBED approach models the same element performances and capabilities throughout the entire integrated SOI's operational environment (i.e., Table 3-4). Not only is each employed DoE designs capable of individually including allocated requirements' unique verification-condition parameters (e.g., $x_1\boldsymbol{B}$, $x_2\boldsymbol{\Gamma}$, and $x_3\boldsymbol{E}$ for elements $\boldsymbol{f_1}$ and $\boldsymbol{f_2}$ as listed in Table 3-3) to sustain RBT&E's existing adherence to the element verification's intrinsic *Requirement-Based Performance Verification* goal. In return, SBED improves the T&E community's ability to comprehensively verify the element function-under-test requirement fulfillment.

Furthermore, if Fig 3-1's four lower-level elements are assumed to collectively achieve their corresponding SOI capability from start to end (e.g., as $\boldsymbol{G}$ from Figure 3-1), Table 3-5's RBT&E element-data conclusions cannot be quantitatively amalgamated to produce the respective SOI capability forecast due to their existing RBT&E's data-inconsistency challenges. However, corresponding SBED's (normalized and standardized) *element-capability* forecasts, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$s, *for i = 1 to 4* from Figure 3-11 to Figure 3-14, are capable of being quantitatively amalgamated to deliver an objective SOI-capability

forecast. Thus, concurrently accomplish the *SOI-Capability Identification* T&E goal as part of SBED Step 5.

Table 3 - 5, Reporting Formats of the Traditional RBT&E Verification Data

Element Requirement #	Test Conditions	Binomial Evaluation Results	Statistical Analysis Results [B] ($\mu \pm 2\cdot\sigma$)	Stochastic Results *# of failures/total sample*
f_1	x_1*B, x_2*Γ, x_3*E (combined)	2 failures (31 total)	(10.56A ±4.1) | x_1*B and x_2*Γ	$P(7.5 \leq f_1 \leq 15\ A)$ | x_1*B *and* x_2*Γ = 93.5% (29 out of 31)
f_2		6 failures (32 total)	(68.8Δ ± 28) | x_1*B and x_3*E	$P(f_2 \leq 80\ \Delta)$ | x_1*B *and* x_3*E = 81%, (26 out of 32)
f_3	$x_{4(min)}Z \leq x_4 \leq x_{4(max)}\ Z$	4 failures (15 total)	(57.1E ±16) |{$x_{4(min)}Z \leq x_4 \leq x_{4(max)}\ Z$}	$P(f_3 \leq 75\ E)$ |{ $x_{4(min)}Z \leq x_4 \leq x_{4(max)}\ Z$) = 73% (11 out of 15)
f_4	x_5*Θ, x_6*I	all pass	(23H ±10.2) | x_5*Θ and x_6*I	$P(f_4 \geq 15\ H)$ | x_5*Θ *and* x_6*I = 100%
SOI Capability		*Inconclusive*	*Inconclusive*	{$p(f_1)$ ***AND*** {$p(f_2)$ ***OR*** $p(f_3)$} ***AND*** $p(f_4)$} | x_1=?, x_2=?, x_3=?, x_4=?, x_5=? = 95%

Note A: All calculations are based on the raw data (as presented in Appendix A) and Table 2-1's arbitrary allocated requirement thresholds.

Note B: Prediction interval of ($2\cdot\sigma$) is arbitrarily selected for this illustrative example.

3.2.5 Step 5 Forecasting the SOI-Capability

Once SBED's RSM analytical techniques from Step 4 are applied to every element's raw performance verification results and have characterized each corresponding element performances as normalized and standardized *element-capability* response models, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$s, these response models are to be used to produce an objective SOI-capability forecast, $\hat{\boldsymbol{F}}_{\boldsymbol{capability}}\boldsymbol{(R)}$, either through: (a) their quantitative amalgamation via Boolean-Algebraic summation to capture the elements' net response, or (b) their aggregate coherent simulation of respective MBSE low-level models. Both SOI-capability forecasting methods are to be discussed in the following subsections.

3.2.5.1 Through Superposition of Element-Capability Models

The first method to obtain the SOI capability forecast, $\hat{\boldsymbol{F}}_{\boldsymbol{capability}}\boldsymbol{(R)}$, is through the quantitative amalgamation of all calculated (and now normalized and standardized) *element-capability* forecasts, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$s. SBED first calculates each R_{x_R} subspace's SOI-capability point estimate, $\hat{F}_{capability}(R_{x_R})$s, through the amalgamation of all corresponding *element-capability* point estimates, $\hat{f}_{i_capability}(R_{x_R})$s. As mathematically described in Eq. 3-22, which combines Eq. 3-2 with either Eq. 3-17 or Eq 3-18 (depending on $\hat{f}_{i_capability}(R_{x_R})$s' format, probability or discrete). The amalgamation of a $\hat{F}_{capability}(R_{x_R})$ point estimate to capture the elements' net response is theoretically visualized as a superposition of all corresponding $\hat{f}_{i_capability}(R_{x_R})$s w.r.t. the sequenced SOI capability scenario diagram (i.e., set $\boldsymbol{G}$ from Eq. 3-1) as introduced in Myers & Montgomery (2002). With interpretation of the De Morgan's Law as presented in Bertsekas

& Tsitsiklis (2002) and presented in Eq. 3-23, this superposition can be achieved through Boolean algebra summation as explained in Eq. 3-24 and Eq. 3-25 for probability and discrete formats of $\hat{f}_{i_capability}(R_{x_R})$s respectively.

$$\hat{F}_{capability}(R_{x_R}) \overset{\text{def}}{\equiv} \boldsymbol{G}(R_{x_R})$$
$$= \{\forall \{\hat{\mathrm{p}}\left(f_i \textbf{\textit{ meeting }} \mathrm{T}_{R_{x_R}} \middle| R_{x_R}\right)\} \in F(R_{x_R})\}, \text{ or}$$
$$= \{ \forall \{\textbf{\textit{meeting }} \mathrm{T}_{R_{x_R}} | R_{x_R}\} \in F(R_{x_R})\} \quad \text{(Eq 3-22)}$$

$$A = \{\bar{A}\}', \text{ and } A \cap B \cap C \cap \cdots = \{\overline{A \cap B \cap C \cap \cdots}\}'$$
$$= \{\bar{A} \cup \bar{B} \cup \bar{C} \cup \cdots\}' \quad \text{(Eq. 3-23)}$$

$$\hat{F}_{capability}(R_{x_R}) = \cap_i^m \hat{p}_i(f_i \textbf{\textit{ meeting }} T_{\boldsymbol{f_i}} | R_{x_R})$$
$$= \cup_i^m \{\overline{\hat{p}_i(f_i \textbf{\textit{ violating }} T_{\boldsymbol{f_i}} | R_{x_R})}\},$$
$$\textit{where } P_i \cup P_j = P_i + P_j - (P_i \cdot P_j). \quad \text{(Eq. 3-24)}$$

$$\hat{F}_{capability}(R_{x_R}) = \cap_i^m (\hat{f}_i \textbf{\textit{ meeting }} T_{\boldsymbol{f_i}} | R_{x_R})$$
$$= \cup_i^m \{\hat{f}_i \textbf{\textit{ violating }} T_{\boldsymbol{f_i}} | R_{x_R}\},$$
$$\textit{where } \cap_i^m \hat{f}_i \textbf{\textit{ meeting }} T_{f_i} \overset{\text{def}}{\equiv} \textit{SOI meeting SOI requirement at } R_{x_R},$$
$$\textit{and } f_1 \textbf{\textit{ violating }} T_{\boldsymbol{f_i}} \cap \{\cap_i^{m-1}(\hat{f}_i \textbf{\textit{ meeting }} T_{f_i} | R_{x_R})\} \overset{\text{def}}{\equiv} \text{SOI } \textit{not meeting} \text{ SOI requirement at } R_{x_R}. \quad \text{(Eq. 3-25)}$$

To briefly demonstrate this super positioning of individual corresponding *element-capability* point estimates through the continuum of the same arbitrary to-be-forecast SOI-

capability example previously from Subsection 3.2.4.5 (Resolution of RBT&E's Data-Inconsistency Challenges), the four *element-capability* models, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$*, for i=1 to 4,* as illustrated from Figure 3-15 are Boolean-algebraically amalgamated to produce their corresponding SOI-capability, $\widehat{\boldsymbol{F}}_{\boldsymbol{capability}}(\boldsymbol{R})$, in probability format. This super positioning process is explained in Eq. 3-26 (as derived from Eq. 3-24) and illustrated (step by step) in Figure 3-16.

$$\hat{F}_{capability}\left(R_{x_R}\right)$$
$$= \hat{f}_{1_capability}\left(R_{x_R}\right) \cap \left(\hat{f}_{2_capability}\left(R_{x_R}\right) \cup \hat{f}_{3_capability}\left(R_{x_R}\right)\right) \cap \hat{f}_{4_capability}\left(R_{x_R}\right), \forall R_{x_R}\text{s}$$

where $\hat{f}_{i_capability}\left(R_{x_R}\right) \cup \hat{f}_{j_capability}\left(R_{x_R}\right) = P_{f_i} + P_{f_j} - \left(P_{f_i} \cdot P_{f_j}\right) | R_{x_R}$, *and*

$$\hat{f}_{i_capability}\left(R_{x_R}\right) \cap \hat{f}_{j_capability}\left(R_{x_R}\right) = \left(P_{f_i} \cdot P_{f_j}\right) | R_{x_R} \qquad \text{(Eq. 3-26)}$$

Finally, the SOI-capability forecast can now be objectively obtained with the consolidation of all $(x_R)^k$ number of calculated SOI-capability point estimates, $\hat{F}_{i_capability}\left(R_{x_R}\right)$s, into the *k*-dimensional arrays of $\boldsymbol{R}$.

3.2.5.2 Inputs to MBSE Models

In addition to the SBED's quantitative amalgamation of *element-capability* response models, there exists another method to produce SOI-capability forecasts through MBSE. SBED-produced *element-capability* models, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$s, can also be delivered and utilized as simulation inputs to respective MBSE lower-level element models. Since these MBSE element models are now represented with data-metrological

and -domain matching variables, an objective SOI mission-representative capability forecast can be aggregately and coherently simulated (Li, et al., 2018b).

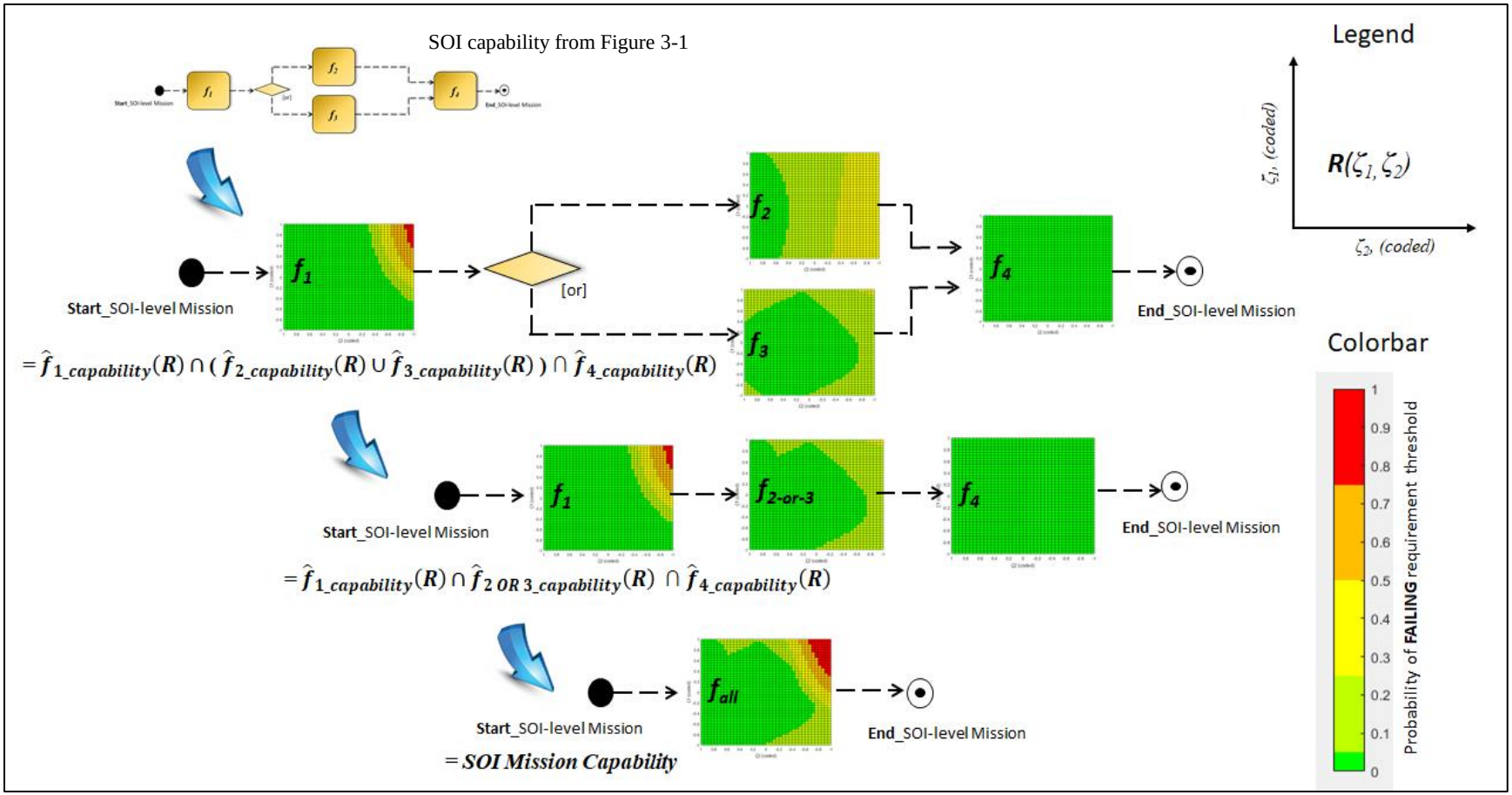


Figure 3 - 16, Amalgamation of SOI-Capability via Boolean-Algebraic Summations

3.3 Benefits of SBED

SBED introduces a top-down SOI-synoptic verification perspective to supplement the traditional RBT&E's existing bottom-up reductionist verification approach during the SOI Implementation process. In addition to preserving all of RBT&E's inherent benefits towards the accomplishment of *Requirement-Based Performance Verification* T&E goal's advantages as previously discussed in Subsection 2.1.2 (Benefits of the Traditional RBT&E Approach), SBED further introduces several additional SE and T&E benefits throughout SOI's development cycle. These SBED-introduced benefits are discussed next and further demonstrated in this dissertation's data-analysis example.

3.3.1 Early Delivery of SOI Capability Forecast Through SBED

The first SBED-introduced benefit to the SOI Implementation process is the early delivery of an objectively SOI-capability forecast, which has been previously unobtainable through the employment of traditional RBT&E test approach (Helle & Schamai, 2016; Fiebrandt & Wilson, 2008; Shaikh & Moore, 1997). Per Li, et al. (2018a), SBED provides a complete solution to the existing RBT&E's data-inconsistency challenges of arithmetical nonuniformity and nonlinearity in variable domains (i.e., mismatching as discussed in Subsection 2.2.2., [Inadequacies of the RBT&E Approach]) and enables the Implement process' concurrent accomplishment of the *SOI-Capability Identification* T&E goal. The delivery of the SBED produced SOI-capability forecast not only objectively informs decision makers on specific section(s) within the SOI operational environment where (and where not) potential SOI failures may exist, it also objectively identifies the culpable child element(s) associated with these forecasted SOI risks and uncertainties.

3.3.2 Enhancement of the Requirement-Based Performance Verification Goal

SBED also introduces direct benefits towards the SOI Implementation process's ability to accomplish the primary *Requirement-Based Performance Verification* T&E goal. With the delivery of SBED's *element-capability* models that comprehensively characterize every respective element performance w.r.t. the entire SOI operational environment, the scope of Implementation process' primary *Requirement-Based Performance Verification* T&E goal (which has been previously not expandable per Streit, et al. [2008]) can be objectively expanded to beyond just element requirement-specified verification conditions (Bjorkman & Zeis, 2007; Eldh, et al., 2010; Hsueh, 1995).

3.3.3 Enablement of MBSE Supportability

Collectively, the delivery of both SBED's *element-capability* models and its SOI-capability forecast regression models also improves the Implementation process' supportability to MBSE's SOI-capability identification by: (a) Resolving the existing data-domain inconsistency between RBT&E and MBSE by providing the ability to coherent simulate an SOI-capability forecast that is comprehensively representative of the intended SOI operational environment (Li, et al., 2108b); and (b) introducing objective means to validate respective low-level models corresponding to aspects of the SOI-design (Karban, et al., 2008). Both of these two MBSE supportability attributes are previously unobtainable through provisions of traditional RBT&E's mismatched verification data.

3.3.4 Preservation of RBT&E's Intrinsic Benefits

As SBED supplements the traditional RBT&E approach with a top-down SOI-synoptic verification perspective, it also preserves RBT&E's benefits from generating isolated test artifacts specific only to the individual element functions-under-test. RBT&E's existing scope, which "advantageously" quarantines RBT&E-collected element verification data from the influences of potential interactions of other element functions, is considered as an intrinsic benefit to the SOI Implementation process and is concomitant to both the establishment of clear and bidirectional traceability between individual element verification results (as well as their identified element performance defects). The supplementation of SBED does not remove this intrinsic benefit. This is because SBED preserves the traditional RBT&E's methodology of executing individual element verification tests as independent and autonomous entities.

3.3.5 Improvements to Remaining RBT&E' Intrinsic Benefits

Unlike the traditional RBT&E test approach which is flexible to SOI design process's nature of constant requirement changes, creeps, and churns, SBED methodology introduces flexibility and adaptability without sacrificing its top-down SOI-synoptic verification perspective. Given that it is unrealistic to repeat the applicable element verification test upon every SOI design change (Eldh, et al., 2010; Dahmann, 2012), SBED's RSM analytical techniques can augment the new element verification results (that correspond to new design changes) and produce updated SOI-capability forecasts via either Boolean-algebraic summations or MBSE coherent simulations. SBED's improvement, of the Implementation process' previous "robustness" towards constant element-design and -

requirement changes, becomes a valuable benefit when it comes to the ever-evolving Complex System design strategy (Hesari, et al., 2013).

With the most common areas of an element-requirement change are introduced in the form of threshold-value adjustment or verification-condition adjustments. For example, reduce the demonstration of element A's y_1-mph speed under x_1°C verification condition to $(y_1 - y^*)$ mph speed now under $(x_1 - x^*)$°C condition. If certain allocated requirement threshold-value(s) has been modified, the only procedural effect to SBED's implementation includes in repeating its Step 4's RSM calculations to recompute *element-capability* regression models and beyond (i.e., during the post-verification test analysis only). While the previous flexibility aspect of SBED is also applicable to the traditional RBT&E test approach, the same cannot be said when it comes to modification of allocated requirement's verification-condition. When there is an adjustment to allocated requirement verification condition, previous RBT&E's verification data (i.e., performance and verification data specifically at the requirement's verification conditions) may no longer be applicable; however, the only procedural effect to SBED's implementation remains in repeating its Step 4's RSM calculations to redetermine *element-capability* regression models (i.e., also only during the post-verification test analysis). This is because SBED's collected element verification data (i.e., through DoE designs) have already fully captured the respective element performances w.r.t. the entire observed domain (or intended operational environment).

Additionally, SBED's methodology also introduces new flexibility and adaptability against changes at the SOI-design level. For example, given the previous definition of a SOI capability is modified with the addition of new element requirement(s) during the

Implementation process. The corresponding SOI-capability forecast can still be either quantitatively amalgamated or coherently simulated via MBSE as long as the respective element verification data captures the same causality relationship w.r.t. the overall SOI's operational environment. The only procedural effect to SBED's implementation includes in the repeat its Step 5's amalgamation of *element-capability* response models to reproduce an updated SOI-capability forecast (i.e., also only during the post-verification test analysis). Similarly, if an existing SOI-capability definition is modified through either a removal or a rearrangement of already interconnected lower-level elements, once again, the only procedural effect to SBED's implementation includes the repeat its Step 5.

3.3.6 Benefits to SE Processes

Lastly, and previously as illustrated in Figure 1-2, the supplementation of SBED's top-down SOI-synoptic verification perspective to traditional RBT&E's bottom-up reductionist element verification approach also establishes the SOI Implementation process' SE-information feedforward to future Integration and Verification processes along the SE-Vee. It provides objective means to both understand the SOI emergent behaviors w.r.t. the SOI hierarchical structure and incorporating element interfaces (Anandalingam & Friesz, 1992) as well as to identify subsequent integrated test subjects (Zhu et al., 2016). Concurrently and per Gotel & Finkelstein (1994), SBED's supplementation also establish the SOI Implementation process's SE-information feedback to previous adjacent SOI design Decomposition and Definition processes along the SE-Vee and provides objective means to understand the existing SOI design and SOI-design products (e.g. MBSE lower-level models).

As the SOI development proceeds to subsequent Integration and Verification processes with higher integrated SOI-level functions-under-test. This same SBED methodology can be iteratively reapplied to continuously provide SE-information feedforward to support the capability identification effort at the next higher integrated levels and provide SE-information feedback to validate the next higher adjacent SOI Decomposition and Definition process' corresponding design models and products. When the SOI-capability forecast of one integrated SOI is grouped with other "sibling" capability forecasts (all w.r.t. the common operational environment domain), SBED is able to continue to objectively forecast the next integrated-level's capability until the ultimate SOI-level's actual integration and validation activities.

3.4 Limitations of the SBED Test Approach

Given the SBED test approach's ability to produce an objective integrated SOI-level capability forecast and to inform decision makers on specific section(s) within the SOI operational environment where (and where not) potential SOI failures may exist, it is to note that this study's intent to present the SBED test approach as a direct substitute for actual SOI integration and V&V activities subsequent to the Implementation process. This is because the developed SBED methodology still includes a significant limitation. The most significant being that it does not introduce the ability to identify or predict actual interaction behaviors amongst the elements within a specific SOI design. During the SOI Implementation process, there remains a very limited understanding of SOI's inter-element behavior and the actual relationships between element-level performances and SOI-level performances (Deplanques, et al., 1996; Jones & Deshmukh, 2005; Fiebrandt & Wilson,

2008). Despite SBED's supplementation of a qualitative and top-down SOI-synoptic verification and analytical verification approach, the absence of actual integration-verification data prohibits any objective interpretation and prediction of inter-element performances and behavior. Therefore, SBED's ability to define (or forecast) SOI capability through standalone analysis of individual element performances cannot be considered as an equivalent replacement to the actual integrated-level V&Vs (Bjorkman, et al., 2005; Loren, 2007).

4 Data Analysis Example

SBED's ability to objectively forecast SOI capability, through the quantitative amalgamation of individual collected element performance and verification data, is demonstrated through a data-analysis examples in this chapter. This example exhibits SBED's ability to resolve RBT&E's data-metrological and –domain inconsistency challenges and to objectively forecast SOI capability (i.e., normalize, standardize, and amalgamate of previously mismatching element verification data) by applying its DoE and RSM analytical techniques onto an existing US Air Force Flight Test Center (AFFTC) T-38C flight test data set as presented in Gunn-Golkin, et al. (2010).

The scope of this data analysis example includes one of AFFTC's original test objectives, to assess the Operator Workload (OWL) mission-suitability of USAF legacy T-38A/B airframes' landing technique (hereon designated as 'L_1') as demonstrated on the then new USAF T-38C airframe (Gunn-Golkin, et al., 2010). Within this example, SBED-introduced SE and T&E benefits are to be exhibited by comparing SBED's quantitatively-produced capability forecast against AFFTC's original qualitative SOI assessment conclusions (as subjectively drawn from standalone analyses of individual RBT&E-collected element verification data).

4.1 AFFTC Test Method Description and Original Conclusions

AFFTC's original L_1 OWL-suitability assessment includes two independent and concurrently executed methods. First, a qualitative interpretation of test pilots' Likert-scale survey ratings collected after every executed landing attempt to holistically characterize L_1's OWL demand at the airframe-system level. The second includes standalone

quantitative comparisons between L_1's collected element-level verification data against the baseline T-38C airframe's landing technique's (hereon designated as 'L_2') equivalents. In reference to Bortolusi, Kantowitz, & Hart's research from 1986 which suggests that OWL-demand differences between two independent tasks can be quantified through comparison between their demonstrated performances variances, AFFTC statistically compares the variability of each of L_1's collected lower-level element verification data against its L_2 equivalents' to determine equality. L_1's OWL demand is deemed as mission suitable if, and only if, its element data variabilities are proven to be statistically equal (or better) to L_2's. A total of four common standalone element verification data (collected from both L_1 and L_2's executed landing attempts during test) are used as part of this quantitative comparison assessment method and they are listed in Table 4-1.

Table 4 - 1, AFFTC Test's Standalone Lower-Level Verification Data

ID	Variables & Units of Measurement	Descriptions
f_1	Touchdown (TD) Distance, ft	Distance of aircraft touchdown location beyond runway threshold.
f_2	TD Speed, KCAS	Aircraft speed at touchdown.
f_3	Height Above Threshold, ft	Aircraft height above the runway threshold.
f_4	Vertical Velocity at TD, ft/min	Vertical component of aircraft speed at touchdown.

Despite using the same executed L_1 landing attempts to collect their respective raw data, two contradicting L_1 OWL-suitability assessments conclusions are drawn from the two AFFTC assessment methods. From the first (qualitative) assessment method, interpretation of the qualitative pilot survey ratings suggests a significant portion of executed L_1 landing attempts demonstrate noticeable increased OWL demand w.r.t. L_2's

landing attempts (Gunn-Golkin, et al., 2010, p. A-12). Breakdown of the pilot survey ratings is shown in Table 4-2.

Table 4 - 2, AFFTC Test's Pilot OWL Survey Ratings

	Very Difficult	Difficult	Average	Easy	Very Easy	Average Rating
	(1)	(2)	(3)	(4)	(5)	
L_1	0	8	18	4	0	**2.87**
L_2	0	0	6	19	5	**3.96**

Conversely from AFFTC's second OWL-suitability assessment method, all the quantitative variability-comparison results between L_1's and L_2's four standalone element verification data (from Table 4-1) fail to reveal any statistically conclusive differences. Thus, they all independently lead to the contradictory conclusion that no significant increase in L_1's OWL-demand (w.r.t. the L_2 baseline) is demonstrated (Gunn-Golkin, et al., 2010, p. 26). Furthermore, since all of L_1 and L_2's quantitative standalone element verification data are reported (and analyzed) in RBT&E's mismatching scalar statistic format like the examples provided in Table 2-2 (Gunn-Golkin, et al., 2010, Appendix C), AFFTC is unable to further quantitatively amalgamate L_1's four element verification data to produce an integrated L_1 OWL-suitability forecast (at the airframe-system level) to objectively confirm or reject the contradicting qualitative survey-rating assessment's conclusion.

This example suggests that if this study's SBED is employed as a supplement to AFFTC test's (RBT&E-collected) L_1 element verification data, a quantitative L_1 airframe-system level OWL-suitability forecast can be objectively produced. It is this data analysis

example's intent to: (a) retroactively apply SBED's DoE and RSM data-analysis techniques on to the existing AFFTC element verification data to independently produce L_1's quantitative OWL-suitability forecast; and (b) compare this SBED-produced forecast against AFFTC's original qualitative survey-rating assessment's conclusion to demonstrate (and validate) SBED's methodology and introduced advantages.

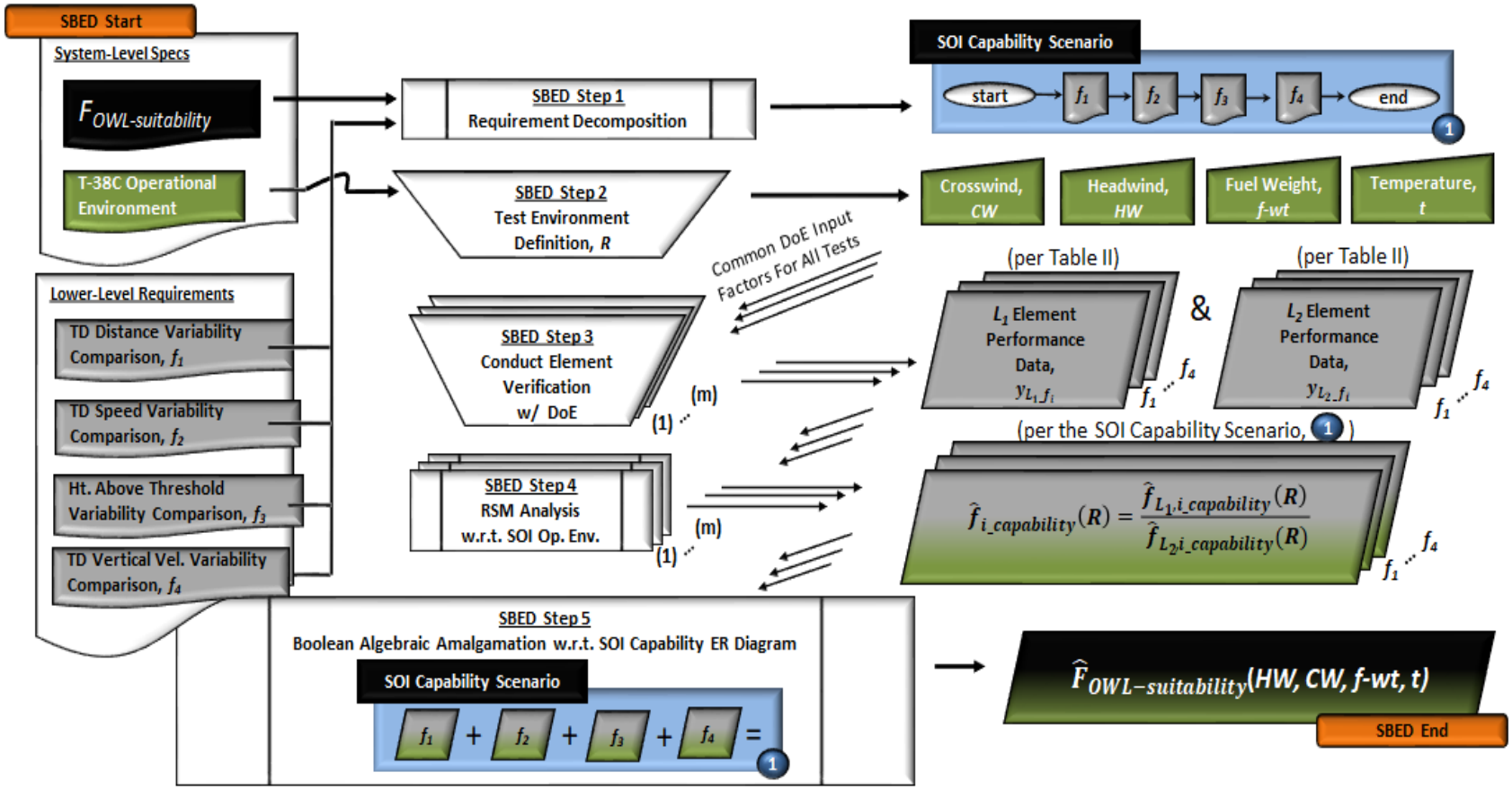


Figure 4 - 1, The SBED Model Diagram and Data Flow Chart as Applied to the Data-Analysis Example

4.2 Overview of the SBED Data Analysis Example

As shown in Figure 4-1 (top left at SBED Start), this example designates L_1's OWL suitability as this example's to-be-forecasted SOI capability, $\boldsymbol{F_{capability} \stackrel{\text{def}}{=} F_{OWL\text{-}suitability}}$; and designate the L_1:L_2 data-variability comparisons of the four standalone lower-level element verification data (from Table 4-1) as $\boldsymbol{F_{OWL\text{-}suitability}}$'s four elements, $\boldsymbol{f_i}$s. This implies that every L_1 and L_2's standalone element verification data is first independently modeled as prerequisite RSM-calculated response models, $\hat{f}_{L_1_i}(R)$s and $\hat{f}_{L_2_i}(R)$s. Then, their respective scalar prediction confidence intervals, $\hat{f}_{L_1_i_variaiblity}(R_{x_R})$s and $\hat{f}_{L_2_i_variaiblity}(R_{x_R})$s, are compared against each other. The ratios are then consolidated to construct $\boldsymbol{F_{OWL\text{-}suitability}}$'s element mean-response models, $\boldsymbol{\hat{f}_i(R) = \hat{f}_{L_1_i_variaiblity}(R) : \hat{f}_{L_2_i_variaiblity}(R)}$, *for i=1,…,4*.

This example also designates ($T_i \leq 1.5$) as $\boldsymbol{F_{OWL\text{-}suitability}}$'s four elements' requirement thresholds. (This T_i corresponds with the rejection-region-of-an-equality of a two-variances F-test based on $\alpha = 0.10$ and AFFTC test's executed L_1 and L_2's sample sizes of 29 and 30 executed landing attempts respectively (Ott & Longnecker, 2015).) For this example, ($T_i \leq 1.5$) implies that every R_{x_R} subspace (within the yet-to-be-defined $\boldsymbol{R}$), whose point estimates, $\hat{f}_i(R_{x_R}) = \hat{f}_{L_1_i_variaiblity}(R_{x_R}) : \hat{f}_{L_2_i_variaiblity}(R_{x_R})$, exceeds above *1.5* (i.e., significant increase in L_1 data variability w.r.t. to the L_2 baseline), is to be considered as a L_1 operating condition in which its OWL-demand may not be mission suitable per Bortolussi, et al. (1986). Concurrently, these non-mission suitable R_{x_R} subspaces' respective *element-capability* point estimates (in discrete format per Eq. 3-18), $\hat{f}_{i_capability}(R_{x_R})^*$s, are to be interpreted as ($\boldsymbol{violating}\ T_i|R_{x_R}$). Knowing that there

exist such non-mission suitable R_{x_R} subspaces, as suggested by AFFTC's actual qualitative survey-ratings' conclusion (from Table 4-2), this study have successfully demonstrated and validated SBED's methodology and benefits if, and only if, the calculated L_1's OWL-suitability forecast, $\widehat{\boldsymbol{F}}_{\boldsymbol{OWL-suitability}}(\boldsymbol{R})$, also contains these non-mission suitable R_{x_R} subspace(s) to corroborate with AFFTC's integrated-level assessment conclusion. This is expressed as this data-analysis example's null hypothesis in Eq. 4-1.

$\boldsymbol{H_0}$: $\hat{f}_{L_1_i_variaiblity}(R_{x_R}) \leq \hat{f}_{L_2_i_variaiblity}(R_{x_R})$, $\forall\ \hat{f}_i(R_{x_R})$s, *and*

$\boldsymbol{H_a}$: $\hat{f}_{L_1_i_variaiblity}(R_{x_R}) > \hat{f}_{L_2_i_variaiblity}(R_{x_R})$, *for at least 1* $\hat{f}_i(R_{x_R})$, *where i = 1, …, 4.*

T.S.: $\hat{f}_{L_1_i_variaiblity}(R_{x_R}) \,/\, \hat{f}_{L_2_i_variaiblity}(R_{x_R})$,

R.R.: Reject $\boldsymbol{H_o}$ if **T.S.** is $\geq$ 1.50 for at least 1 $\hat{f}_i(R_{x_R})$. (Eq. 4-1)

4.3 The SBED Data Analysis of the T-38 Test Data

4.3.1 Step 1 Develop a Scenario-Based $\mathbf{F_{OWL-suitabiltiy}}$ Definition

A scenario diagram depicting $\boldsymbol{F_{OWL-suitabiltiy}}$ is created and shown in Figure 4-1 (top right, Step 1). In attempt to maintain alignment with AFFTC's original quantitative L_1 OWL-suitability assessment method, this scenario diagram sequences the four elements (from Table 4-2), $\boldsymbol{f_i(R)}=\boldsymbol{f}_{\boldsymbol{L_1},\boldsymbol{i}_{\boldsymbol{variability}}}(\boldsymbol{R})$: $\boldsymbol{f}_{\boldsymbol{L_2},\boldsymbol{i}_{\boldsymbol{variability}}}(\boldsymbol{R})$, for *i=1, …, 4*, in series; and in synchrony with Eq. 3-1, $\boldsymbol{F_{OWL-suitabiltiy}}$'s scenario diagram is mathematically expressed in Eq. 4-2.

$$\boldsymbol{F_{OWL-suitabiltiy}} \coloneqq \mathbf{G}\colon \{\forall \boldsymbol{f_i} \rightarrow \boldsymbol{F_{OWL-suitabiltiy}}\}],\ i = 1 \text{ to } 4$$
$$= \mathbf{G}[\boldsymbol{f_{L_1,i_variability}} \colon \boldsymbol{f_{L_2,i_variability}}] \qquad \text{(Eq. 4-2)}$$

4.3.2 Step 2 Introduce the Common Domain

AFFTC's original element verification data include four environmental and conditional parameters as captured during every L_1 and L_2's executed landing attempts. They are listed in Gunn-Golkin, et al. (2010), Appendix C and in Table 4-3. Even though these captured parameters are not utilized (nor needed) as part of AFFTC's original RBT&E data analysis, they are to be utilized as this example's SBED common domain, $\boldsymbol{R}(\xi_j\text{s})$. As shown in Eq. 4-3 and in synchrony to Eq. 3-2, all subsequent *element-capability* forecasts are to be normalized w.r.t. Table 4-3's $\boldsymbol{R}(\xi_j\text{s})$ that is visually illustrated as a (k=3)-dimensional $\boldsymbol{R}$ space (CW_x by HW_y by $f\text{-}wt_z$) of which the temperature parameter, t, is kept incrementally constant as the $k=4^{th}$ dimension.

$$\boldsymbol{F_{OWL-suitability}} \coloneqq \boldsymbol{G}(\boldsymbol{R}) = \boldsymbol{G}\left(\xi_{CW}, \xi_{HW}, \xi_{f_{wt}}, \xi_t\right)$$
$$= \left\{\forall \mathbf{f}_{\mathrm{i}}\left(\xi_{\mathrm{CW}}, \xi_{\mathrm{HW}}, \xi_{\mathrm{f_{wt}}}, \xi_{\mathrm{t}}\right) \rightarrow \mathbf{F}_{\mathrm{OWL}}\right\} \qquad \text{(Eq. 4-3)}$$

Table 4 - 3, T38C Test's Observation Region

Parameters	designations	Units	Interval (actual)	Interval (coded)
Crosswind Component[1]	*CW*	Knots	{-6.0, 6.8}	{-1, 1}
Headwind Component[1]	*HW*	Knots	{-5.0, 14.7}	
A/C Fuel Weight	*fwt*	Pounds	{1.5K, 4K}	
Ambient Temperature	*t*	^{o}C	{0, 22}	

Note: indicate parameters partitioned from the collected wind-direction and wind-velocity data w.r.t. the landing attempts' runways.

To optimize the *element-capability* model's display resolution, the example selects value of $x_R = 41$ (or 404 number of R_{x_R} subspaces within $\boldsymbol{R}$) to calculate and illustrate its *element-capability* forecasts. The corresponding volume of each R_{x_R} subspace, as related to $x_R = 41$, is defined in Table 4-4.

Table 4 - 4, T-38C Operational Environment Data-Resolution Definitions

Designation	Interval Length	Designation	Interval Length
CW	0.64 Knots	*f-wt*	95 Lbs.
HW	1.04 Knots	*t*	1.1 C^{o}

Note: In coded variables, each X_R represents a R_{x_R} subspace size of 0.05[4]

4.3.3 Step 3 Design and Collect Element Verification Data with DoE

During the AFFTC test, L_1's standalone lower-level element verification data (from both Table 4-1) are captured from 29 executed landing attempts, and L_2's data are from 30 (Gunn-Golkin, et al., 2010, Appendix C). Even though the AFFTC test does not employ (nor needs to employ) any specific DoE designs during its RBT&E data collections and analyses, because these 59 L_1 and L_2 landing attempts are executed at various operating conditions throughout the test's observation region and at a randomized order (as shown in Figure 4-2), employment of a random DoE design can be assumed for this example. Therefore, the raw DoE element verification data collected during both L_1 and L_2's landing

attempts, $\boldsymbol{y_{L_1_f_i}}$s and $\boldsymbol{y_{L_2_f_i}}$s as listed in Gunn-Golkin, et al. (2010), Appendix C, are to be used as inputs to calculate the prerequisite response models, $\hat{\boldsymbol{f}}_{\boldsymbol{L_1_i}}(\boldsymbol{R})$s and $\hat{\boldsymbol{f}}_{\boldsymbol{L_2_i}}(\boldsymbol{R})$s, *for i=1,…, 4*. They are then to be subsequently compared to produce $\boldsymbol{F_{OWL\text{-}suitability}}$'s four lower-level elements' capability forecasts, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})= \hat{\boldsymbol{f}}_{\boldsymbol{L_1_i}}(\boldsymbol{R}): \hat{\boldsymbol{f}}_{\boldsymbol{L_2_i}}(\boldsymbol{R})$, in Step 4.

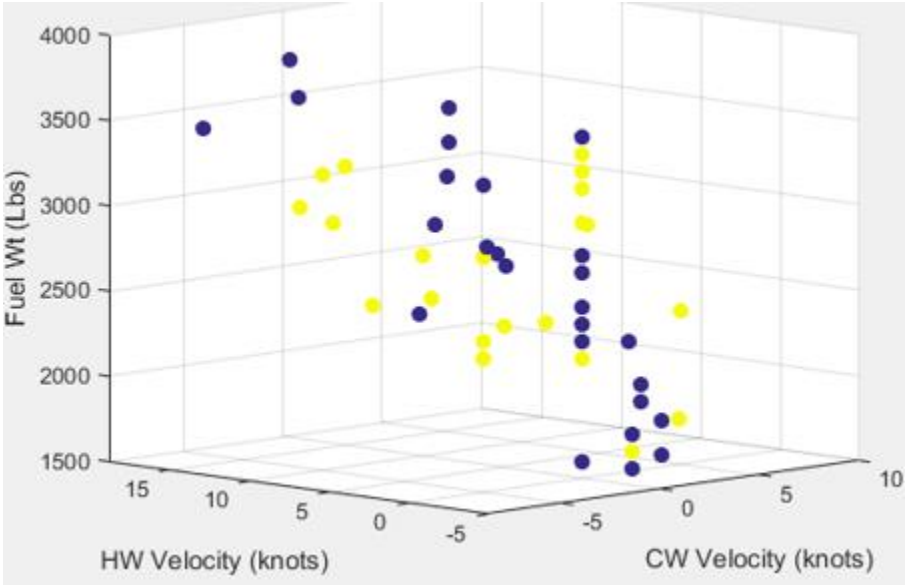


Figure 4 - 2, AFFTC Data-Collection (DoE Design Points) Locations

Figure 4-2. Purple indicates L_1 attempts and yellow indicates L_2's.

4.3.4 Step 4 Calculate *Element-Capability* Forecasts

Through application of RSM's least square method from Eq. 3-7 to Eq. 3-11 and additional stepwise- or backwards stepwise- elimination regression methods (Myers & Montgomery, 2002; Neter et al., 1996), the prerequisite linear regression models of both L_1 and L_2's four standalone elements verification data, $\hat{\boldsymbol{f}}_{\boldsymbol{L_1_i}}(\boldsymbol{R})$s and $\hat{\boldsymbol{f}}_{\boldsymbol{L_2_i}}(\boldsymbol{R})$s, *for i=1,…,4*, are calculated. These regression models are presented in Table 4-5 as significant

regression coefficients of first- or second-order polynomials from Eq. 3-6, along with their associated R^2 statistics (Ott & Longnecker, 2015).

Subsequently, ratios of $\hat{\boldsymbol{f}}_{\boldsymbol{L_1_i}}(\boldsymbol{R})$s and $\hat{\boldsymbol{f}}_{\boldsymbol{L_2_i}}(\boldsymbol{R})$s' respective scalar prediction confidence intervals at every R_{x_R} subspaces throughout $\boldsymbol{R}$, $\hat{f}_{L_1_i_variaiblity}(R_{x_R})$: $\hat{f}_{L_2_i_variaiblity}(R_{x_R})$, are also calculated and designated as SBED's four element-level mean response point estimates, $\hat{f}_i(R_{x_R})$s. Per Eq. 3-18, all four element mean-response point estimates, $\hat{f}_i(R_{x_R})$s, are individually compared against ($T_i \leq 1.50$) to produce the respective *element-capability* point estimates, $\hat{f}_{i_capability}(R_{x_R})$s, in discrete format at every R_{x_R} subspace throughout $\boldsymbol{R}$. Finally, all R_{x_R} subspaces whose calculated $\hat{f}_{L_1_i_variaiblity}(R_{x_R})$: $\hat{f}_{L_2_i_variaiblity}(R_{x_R})$ exceeds (T_i of *1.5*), or $\hat{f}_{i_capability}(R_{x_R}) \stackrel{\text{def}}{=}$ ***violating*** T_i, are interpreted as non-mission suitable L_1 OWL-demand operating conditions (i.e., significant increase in L_1 OWL-demand w.r.t. to L_2 baseline in accordance to Bortolussi, et al. (1986). These non-mission suitable R_{x_R} subspaces' ratios of $\hat{f}_{L_1_i_variaiblity}(R_{x_R})$: $\hat{f}_{L_2_i_variaiblity}(R_{x_R})$ are plotted in Figure 4-3.

Table 4 - 5, Linear Regression Model Coefficients for Element Response Models

	f_1		f_2		f_3		f_4	
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
Constant	841.9	907.7	1.361	2.202	1.658	3.872	-0.639	-1.434
CW	--	--	-3.157	-2.05	--	--	1.19	0.758
HW	--	--	1.439	--	-0.555	1.682	0.804	--
fwt	492	343.8	-5.493	-4.032	--	--	1.058	0.447
t	--	--	--	--	1.099	0.12	--	-0.783
CW²	--	--	--	--	--	--	-1.497	--
HW²	--	--	--	--	6.38	--	--	--
fwt²	--	--	--	--	--	--	-1.351	--
t²	--	--	--	--	-1.935	1.098	--	--
CW · HW	--	--	2.39	--	--	--	-0.829	--

	f_1		f_2		f_3		f_4	
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CW · fwt	--	--	--	--	--	--	-1.999	--
CW · t	--	--	--	--	--	--	--	2.19
HW · fwt	--	--	--	--	--	--	--	--
HW · t	--	--	--	--	2.744	--	--	--
f wt · t	--	--	--	--	--	--	--	1.542
R^2 (%)	50.88	62.97	75.02	50.43	50.28	47.64	72.93	74.25
R^2_{adj} (%)	48.65	61.28	70.48	46.12	38.98	40.5	63.45	67.47

Note: '--' regression coefficients are assigned to non-significant factors that are removed from the linear regression model through the RSM stepwise technique (Myers & Montgomery, 2002, pp 751–753; Neter et al., 1996).

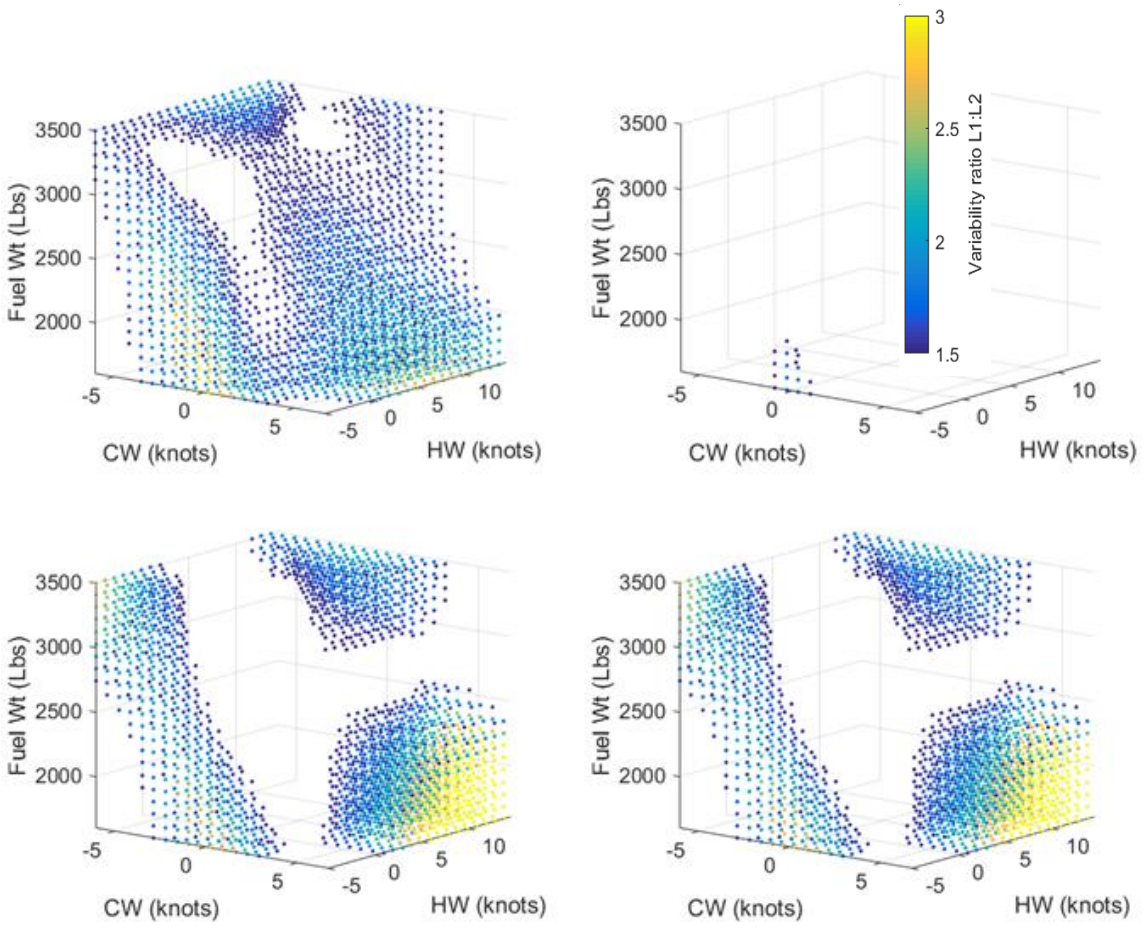


Figure 4 - 3, SBED-Produced T-38C *Element-Capability* Models

Figure 4-3, Color scaled display of forecasted non-mission suitable $\boldsymbol{R_{x_R}}$ subspaces (where L_1:L_2 element data variability ratios exceed 1.5) for SBED elements f_1 (top left), f_2 (top right), f_3 (bottom left), and f_4 (bottom right).

4.3.5 Step 5 Qualitative Amalgamation of SOI-Level Capability Forecast

Per SBED's quantitative amalgamation method to objectively forecast SOI-capability as introduced in Eq. 3-24, L_1's OWL-suitability forecast at the airframe-system level, $\widehat{\boldsymbol{F}}_{\boldsymbol{OWL-suitability}}(\boldsymbol{R})$, can be obtained through the superpositioning of its four calculated *element-capability* forecasts, $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$s. The Boolean-algebraic summation process of $\widehat{\boldsymbol{F}}_{\boldsymbol{OWL-suitability}}(\boldsymbol{R})$'s point estimates, $\hat{F}_{OWL-suitability}(R_{x_R})$s, are performed w.r.t. the $\boldsymbol{F}_{\boldsymbol{OWL-suitabiltiy}}$ scenario diagram in Figure 4-1 (Step 5) and explained in Eq. 4-4 through interpretations of the De Morgan's Law from Eq. 3-23.

$$\hat{F}_{OWL-suitability}(R_{x_R}) = \cap_{\boldsymbol{i=1}}^{\boldsymbol{4}} \hat{f}_{i_{capability}}(R_{x_R})$$
$$= \hat{f}_{1_{capability}}(R_{x_R}) \cap \hat{f}_{2_{capability}}(R_{x_R}) \cap \hat{f}_{3_{capability}}(R_{x_R}) \cap \hat{f}_{4_{capability}}(R_{x_R}) \quad \text{(Eq. 4-4)}$$

The data-analysis example concludes with the consolidation of $(x_R)^k$ number of $\hat{F}_{OWL-suitability}(R_{x_R})$s to construct L_1's airframe system-level OWL-suitability forecast, $\widehat{\boldsymbol{F}}_{\boldsymbol{OWL-suitability}}(\boldsymbol{R})$. It is shown in Figure 4-4.

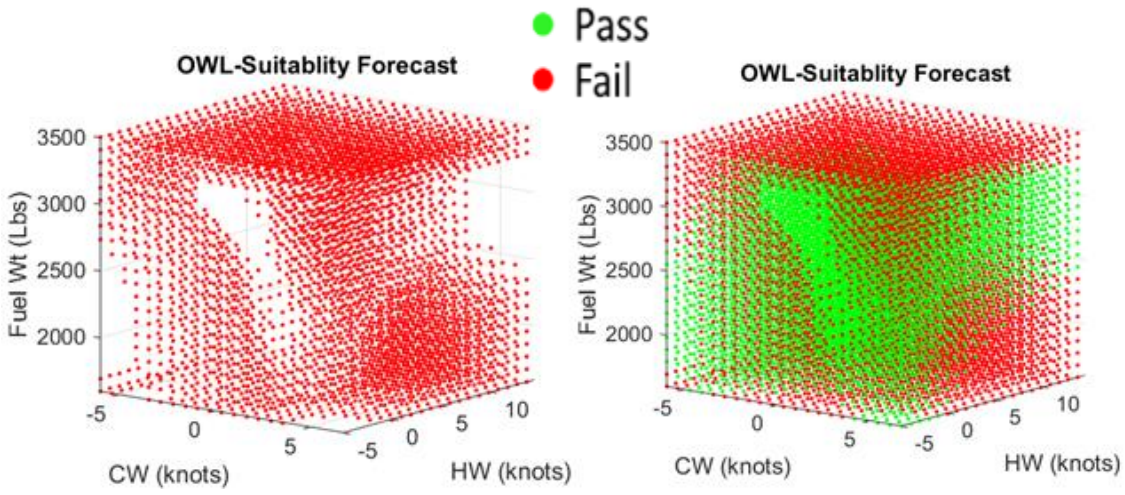


Figure 4 - 4, SBED-Produced T-38 SOI-Level Capability Forecast

Figure 4 – 4, L_1's OWL-suitability forecast, $\widehat{\boldsymbol{F}}_{\boldsymbol{OWL-suitability}}(\boldsymbol{R})$ with non-mission suitable $\boldsymbol{R}_{\boldsymbol{x_R}}$ subspaces shown in red and mission suitable $\boldsymbol{R}_{\boldsymbol{x_R}}$ subspaces in green. Both prediction models' data resolutions are reduced to x_R = 21 (vice 41) for ease of viewing

Figure 4-4's quantitative L_1 OWL-suitability capability forecast at the airframe-system level, previously unobtainable through AFFTC's original RBT&E data and analyses, corroborates with AFFTC's qualitative survey-rating assessment's conclusion (from Table 4-2) which states that there exist multiple sections of non-mission suitable R_{x_R} subspaces within the T-38C operational environment where significant L_1 OWL-demand increases are expected. Specifically, 28,896 out of 68,921 total subspaces, or 41.9% of T-38C's operational environment. With the presence of these forecasted non-mission suitable R_{x_R} subspaces or operating conditions within the T-38C operational enviornment, the authors conclusively reject this data-analysis example's null hypothesis claim from Eq. 4-1 and successfully demonstrate the SBED's ability to obtain normalized and standardized element verification data and quantitatively produce an early SOI capability forecast. Furthermore, this example has exhibited SBED's additional advantages towards expanding

the scope of the primary *Requirement-Based Performance Verification* T&E goal by providing objective evidence of potential risks and uncertainties associated to the SOI design. The SBED-produced $\widehat{\boldsymbol{F}}_{\boldsymbol{OWL-suitability}}(\boldsymbol{R})$ and $\hat{\boldsymbol{f}}_{\boldsymbol{i_capability}}(\boldsymbol{R})$s (Figure 4-3 and Figure 4-4) objectively report that these non-mission suitable L_1 operating conditions correspond with the presence of either heavy fuel weight, or strong tailwind, or combination of both strong headwind and strong crosswind (from the aircraft's starboard direction).

4.4 Validation of the Conclusions

There exists an independent external source to further validate this data-analysis example's L_1 airframe system-level OWL-suitability forecast, $\widehat{\boldsymbol{F}}_{\boldsymbol{OWL-suitability}}(\boldsymbol{R})$. It is the USAF Aircraft Incident Investigation Board Report pertaining to the only USAF T-38C's platform mishap during a final approach on February 2011 (Meidel, 2011). The subject incident involves a pilot attempting to land a T-38C aircraft at the end of a four-solo-sorties day while exercising the same L_1 landing technique as the AFFTC test (Meidel, 2011). During that day's final landing attempt, the aircrew inflicted significant airframe damage due to fatigue, geographical disorientation, and poor in-flight decision making from misinterpretation of instrumentation (Meidel, 2011).

The subject incident can be utilized as a validation source to this example's $\widehat{\boldsymbol{F}}_{\boldsymbol{OWL-suitability}}(\boldsymbol{R})$ because its ambient condition, $\boldsymbol{R}_{\boldsymbol{x_R}}$* (as listed in Table 4-6) is within this example's and AFFTC's test observation region, $\boldsymbol{R}$, (Meidel, 2011). Therefore, and per Eq. 4-4, $\boldsymbol{R}_{\boldsymbol{x_R}}$* subspace's respective scalar L_1 OWL-suitability point estimate, $\widehat{\boldsymbol{F}}_{\boldsymbol{capbility}}(\boldsymbol{R}_{\boldsymbol{x_R}}^{*})$, can be interpolated through this example's Boolean algebraic summation

point-estimate process. It is shown in Table 4-7. With $\hat{f}_{L_1_4_variaiblity}(R_{x_R}{}^*)$: $\hat{f}_{L_2_4_variaiblity}(R_{x_R}{}^*)$ = *1.9*, which is greater than the designated threshold of ($T_i \leq 1.50$), the subject T-38C incident's respective $\hat{F}_{OWL-suitability}(R_{x_R}{}^*)$ is to be interpreted as (violating T_i) per Eq. 3-18 and Eq. 4-1. This implies that, based on the earlier 2010 AFFTC test's forecasted $\hat{F}_{OWL-suitability}(R)$, the interpolation of subject T-38C incident's R_{x_R}* is to be considered as a non-mission suitable L_1 operational condition that may potentially introduce a significant increase in pilot OWL-demand. With this validation, and as suggested by the subject T-38C incident itself, the authors can claim an additional successful validation of this data-analysis example's SBED-produced $\hat{F}_{OWL-suitability}(R_{x_R}{}^*)$ through the following adjudication: When compounded with other unfavorable factors (e.g., fatigue, geographical disorientation, etc.), the presence of increased pilot-OWL demand at the subject T-38C incident's operating condition (as predicted through SBED analysis of AFFTC's 2010 flight test data) introduces the potential risk of reducing pilot's instrumentation mediational awareness and incurring failures of greater magnitude.

Table 4 - 6, USAF FEB 2011T-38C Incident's Ambient Condition

CW	*HW*	*fw-t*	*t*
2.3 (port) knots	-1.9. knots	~1700 Lbs	5° C

Table 4 - 7, Interpolated L_1's OWL-Suitability Point Estimate

	$f_1(R_{x_R}{}^*)$	$f_2(R_{x_R}{}^*)$	$f_3(R_{x_R}{}^*)$	$f_4(R_{x_R}{}^*)$	=	$\hat{F}_{OWL}(R_{x_R}{}^*)$
L_1: L_2 | $R_{x_R}{}^*$	1.37	0.76	1.22	**1.9**	=	***Violate*** (Per Eq. 4-4)
$\hat{f}_{i_{capbility}}(R_{x_R}{}^*)$	*meet*	*meet*	*meet*	***violate***		

5 Discussions, Conclusions, and Future Works

5.1 Discussions and Conclusions

Given the importance to concurrently accomplish the *SOI-Capability Identification* T&E goal throughout all SOI V&V activities, not much emphasis has been devoted during SOI Implementation process' element verification activities due to RBT&E approach's existing data-metrological and –domain inconsistency challenges. This study develops the SBED test approach as a supplement to the traditional RBT&E test approach and provides the ability to normalize and standardize all independently collected element performance verification data in support of their subsequent quantitative amalgamation to objectively forecast SOI capability.

The supplementation of SBED's top-down SOI-synoptic verification approach contributes to both SE and T&E through the introduction of two primary benefits that have been previously unobtainable during the SOI Implementation process. The first benefit is the early delivery of an objective and quantitative SOI capability forecast to support upfront identification and management of SOI design-related risks and uncertainties. The second benefit is the delivery of more-descriptive *element-capability* response models which comprehensively characterize the respective element performances throughout the entire SOI operational environment. Thus, expanding the scope of element performance verifications to beyond just allocated requirement-specified conditions.

5.2 Limitations and Future Research Opportunities

5.2.1 Limitations and Boundaries of Study

As emphasized in Section 3.4 (SBED's Shortcomings), it is not this study's intent to present the SBED test approach as a direct substitute for actual SOI integration and V&V activities subsequent to the Implementation process. Despite SBED's supplementation of a qualitative and top-down SOI-synoptic verification and analytical verification approach, there remains very limited understanding of the SOI's inter-element behavior and the actual relationships between element-level performances and SOI-level performances during the Implementation process (Deplanques, et al., 1996; Jones & Deshmuck, 2005; Fiebrandt & Wilson, 2008). Therefore, SBED's limitations of define (or forecast) SOI capability through only standalone analysis of individual element performances cannot be considered as an equivalent replacement to the actual integrated-level V&Vs (Benjamin, et al., 1999; Loren, 2007).

5.2.2 Recommendations for Future

As discussed in Subsection 3.3.6 (Benefits to SE Process), given the modular nature of a system's hierarchy, SBED can be iteratively repeated throughout the entire SOI V&V processes up until the final SOI-level integrated validation. This scalability advantage of SBED cannot be demonstrated as part of this dissertation because its data analysis example's real-life USAF data set does not include other T-38C airframe-level performance data for subsequent collective amalgamation to produce the next integrated-level capability forecast. Future work is needed to obtain such a multi-hierarchical

integration performance data set and to demonstrate and quantify SBED's scalability at the higher-level integration and verification activities.

Given SBED test approach's success hinges on the identification of common SOI-level operational environment parameters as the common data-domain set of normalize all captured element verification data (in SBED Step 2), inability to determine such common parameter(s) would lead to an immediate halt in SBED's ability to further produce a corresponding SOI-capability forecast. Additional research is needed to identify alternate SOI mission-representative domain set(s) to normalize those independent element verification data that do not share a common integrated operational environment (e.g., System of Systems or SoS). Other data domain parameters, that also bare traits of commonality amongst the dispersed SOI elements (e.g., SoS's constituent systems), need to be selected through other integrated-level sources, purposes, and/or motivations to expand the applicability of SBED methodology.

Implementation of appropriate DoE test methods is a critical component the SBED test approach. However, knowing the real-life challenges associated towards implementing DoE test design during actual element verifications, future research is needed to address this ubiquitous parsimony towards DoE. There is a need to develop another SE- and model-based test frameworks and approaches that can equally support SBED's top-down SOI-synoptic verification approach and can duplicate SBED's resolution of RBT&E test approach's existing data-inconsistency challenges, but without placing a heavy reliance on the implementation of the DoE test methods.

Last and as discussed in Section 1.6 (Scope and Limitations), the propose SBED test approach places a minimal emphasis on the schedule and resource aspects of T&E, as

well as the overall SOI development life cycle. As a result, the delivery of SBED-produced SOI- and element-level capability forecasts cannot solely dictate the outcome of SOI designers and decision maker's responses to immediately implement any respective corrective actions as other SOI development's programmatic constraints are given equal or greater amounts of considerations (i.e., schedule, cost, and resources). Therefore, there remains a need to development and instill a similar SOI-centric T&E approach during the earlier SOI Decomposition and Definition process to elevate the significance of objective integrated-level capability forecasts as future works.

Bibliography/Reference List

Anandalingam, G., & Friesz, T. L., "Hierarchical optimization: an introduction," *Annals of Operations Research*, vol. 34, issue 1, pp. 1-11, 1992, [on line]. Available at: https://www.researchgate.net/profile/Terry_Friesz/publication/226065526_Hierarchical_optimization_An_introduction/links/58fa7fe6aca2723d79d6255f/Hierarchical-optimization-An-introduction.pdf

Apostolov, A., & Vandiver, B., "End to end testing – what should you know?" in *Proceedings of 67th Annual Conference for Relay Engineers,* 2014, pp. 125-131, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/6798999/

Bai, X., Tsai, W. T., Paul, R., Shen, T., & Li, B., "Distributed end-toend testing management," in *Proceedings of 5th IEEE International Enterprise Distributed Object Computing Conference,* 2001, pp. 140-151, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/950430/

Bateson, W., & Mendel, G., *Mendel's principles of heredity*, Cambridge, UK; University Press, 2013, [on line]. Available at: https://scholar.google.com/scholar?hl=en&as_sdt=0,21&q=Bateson+%2B+Mendel+%2B+2013

Benjamin, M., Geist, D., Harman, A., Mas, G., Smeets, R., & Wolfsthal, Y., "A study in coverage-driven test generation," in *Proceedings of the 36th Annual ACM/IEEE Design Automation Conference*, 1999, pp. 970-975, [on line]. Available at: https://dl.acm.org/citation.cfm?id=310108

Berger, C.R., Knowlton, S. W., & Abrahams, M. F., "The hierarchy principle in strategic communication," *Communication Theory,* vol 6, issue 2, pp. 111-142, 1996, [on line]. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-2885.1996.tb00123.x

Bertsekas, D. P., & Tsitsiklis, J. N., *Introduction to Probability, 2nd edition*, Belmont, MA, USA; Athena Scientific, 2008.

Bjarnason, E., Runeson, P., Borg, M., Unterkalmsteiner, M., Engström, E., Regnell B., Sabaliauska, G., Loconsole, A., Gorschek. T., & Feldt, R., "Challenges and practices in aligning requirements with verification and validation: a case study of six companies," *Empirical Software Engineering*, vol 19, issue 6, pp. 1809-1855, 2014, [on line]. Available at: https://link.springer.com/article/10.1007/s10664-013-9263-y

Bjorkman, E. A., & Gray, F. B., "Testing in a joint environment 2004-2008: Findings, conclusions, and recommendations from the joint test and evaluation methodology project," 2009, [on line]. Available at: http://www.dtic.mil/docs/citations/ADA513912.

Bjorkman, E. A., Gray, F. B., March, M. W., "Determining requirements for testing in a joint environment," in *US Air Force T&E Days,* 2005, pp. 1-9.

Bjorkman, E. A., Sarkani, S., & Mazzuchi, T. A., "Using Model-Based Systems Engineering as a Framework for Improving Test & Evaluation Activities," Systems Engineering, vol 16, issue 3, pp. 346-362, 2013, [on line]. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/sys.21241

Bjorkman, E. A., Sarkani, S., & Mazzuchi, T. A., "Test and evaluation resource allocation using uncertainty reduction," *IEEE Transactions on Engineering Management*, vol 60, issue 3, pp. 541-551, 2013, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/6392237/

Bjorkman, E. A., & Zeis, J., "Testing 2020: evaluating weapons systems and capabilities in a joing environment," in *U.S. Air Force T&E Days*, 2007, pp. 1-7. [on line]. Available at: https://arc.aiaa.org/doi/abs/10.2514/6.2007-1602

Bortolusi, M. R., Kantowitz, B. H., & Hart, S. G., "Measuring pilot workload in a motion based trainer: a comparison of four techniques," *Applied Ergonomics*, vol 17, issue 4, pp. 278-283, 1986, [on line]. Available at: https://www.sciencedirect.com/science/article/pii/0003687086901304

Boumen, R., Ruan, S., de Jong, I. S., van de Mortel-Fronczak, J. M., Rooda, J.e., & Pattipati, K. R., "hierarchical test sequencing for complex systems," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol 39, issue 3, pp. 640-649, 2009, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/4787094/

Box, G. E., & Behnken, D. W., "Some new three level designs for the study of quantitative variables," *Technometrics*, vol 2, no 4, pp. 455-475, 1960, [on line]. Available at: https://www.tandfonline.com/doi/abs/10.1080/00401706.1960.10489912

Box, G. E. P., & Draper, N. R., *Response Surfaces, Mixtures, and Ridge Analysis, 2nd edition*, Wiley, 2007.

Box, G. E., & Hunter, J. S., "Multi-factor experimental designs for exploring response surfaces," *The Annals of Mathematical Statistics*, vol 28, issue 2, pp. 195-241, 1957, [on line]. Available at: https://projecteuclid.org/download/pdf_1/euclid.aoms/1177707047

Box, G. E., Hunter, J. S., & Hunger, W. G., *Statistics for Experiments, 2nd edition*, United States: Wiley, 2005.

Box, G. E., & Myer, R. D., "An analysis for unreplicated fractional factorials," *Technometrics*, vol 28, issue 1, pp. 11-18, 1986, [on line]. Available at: https://amstat.tandfonline.com/doi/abs/10.1080/00401706.1986.10488093

Chen, T. Y., Huang, Y. C., Chou, T. S., Shih, C. S., & Liu, J. W., "Model-based development of user-centric automation and assistive devices/systems," *IEEE Systems Journal*, vol 6, issue 3, pp. 338-400, 2012, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/6080700/

Cheng, C. S., "Projection properties of factorial designs for factor screening," in *Screening*, pp. 156-168, New York, NY, USA; Springer, 2006

Chernoff, H., "Locally optimal designs for estimating parameters," *The Annals of Mathematical Statistics*, pp. 586-602, 1953, [on line]. Available at: https://projecteuclid.org/download/pdf_1/euclid.aoms/1177728915

Chipman, H., "Bayesian variable selection with related predictors," *Canadian Journal of Statistics*, vol 24, issue 1, pp. 17-36, 1996, [on line]. Available at: https://arxiv.org/pdf/bayes-an/9510001.pdf, [on line]. Available at: https://arxiv.org/pdf/bayes-an/9510001.pdf

Corral, D., "Requirements engienering needs total systems engineering," *Requirements Engineering*, vol 2, issue 4, pp. 217-219, 1997, [on line]. Available at: https://link.springer.com/article/10.1007/BF02745373

Crisp, H. E., "Systems engineering vision, 2020," Retrieved from: https://www.incose.org/docs/default-source/aboutse/se-vision-2025.pdf , 2015

Dahmann, J., "Integrating systems engineering and test & evaluation in system of systems development," in *Proceedings of IEEE International Systems Conference*, 2012, pp. 1-7, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/6189468/

Dahmann, J., "The state of systems of systems engineering knowledge source," in *Proceedings of 10th System of Systems Engineering Conference (SoSE)*, 2015, pp. 475-479, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/7151979/

Dahmann, J., Lane, J. A., Rebovich, G., & Lowry, R., "Systems of systems test and evaluation challenges," in *Proceedings of 5th International Conference on System of Systems Engineering*, 2010, pp. 1-6, [on line]. Available at: https://www.acq.osd.mil/se/docs/Dahmann-IEEE-SoSE%202010.pdf

Davis, P. K., & Bigelow, J. H., "Motivated metamodels: synthesis of cause-effect reasoning and statistical metamodeling," No. RAND/MR-1570, 2003, Rand Corp., Santa Monica, CA, [on line]. Available at: http://www.dtic.mil/dtic/tr/fulltext/u2/a411888.pdf

de Santiago Junior, V. A. & Vijakumar, N. L., "Generating model-based test cases from natural language requirements for space application software," *Software Quality Journal*, vol 20, issue 1, pp. 77-143, 2012, [on line]. Available at: https://dl.acm.org/citation.cfm?id=2124531

Deaconu, S., & Coleman, H. W., "Limitations of statistical design of experiments approaches in engienering testing," *Journal of Fluids Engienering*, vol 122, issue 2, pp. 254-259, 2000, [on line]. Available at: fluidsengineering.asmedigitalcollection.asme.org/article.aspx?articleid=1429159

Department of Defense (United States), *Joint Capability Integration and Development Systems*, CJCS Instruction 3170.01C, 2003. [on line]. Available at: https://www.dau.mil/acquipedia/Pages/ArticleDetails.aspx?aid=12227505-ba29-41c0-88f0-682a219d5bbc

Department of Defense (United States), *Quadrennial Defense Review Report*, 2010, [on line]. Available at: https://www.defense.gov/Portals/1/features/defenseReviews/QDR/QDR_as_of_29JAN10_1600.pdf

Department of Defense (United States), *Universal Joint Task Manual*, 2011, [on line]. Available at: http://www.jcs.mil/Portals/36/Documents/Library/Manuals/m350004.pdf?ver=2016-02-05-175706-460

Department of Defense (United States), *Defense Acquisition Guidebook*, 2017, [on line]. Available at: https://www.dau.mil.tools.dag.

Department of Navy (United States), Advance Undersea Systems Program Manager, "Concept of operations for the submarine rescue diving recompression system," (Dec)2009.

Department of Navy (United States), Naval Air Systems Command, "Test reporting handbook by naval air systems command, integrated systems evaluation, experimentation and test (ISEET)," 2011.

Deplanques, P., Yriarte, L., Zapata, R., & Sallantin, J., "A methodology for complex systems test and validation," in *Proceedings of international Conference on Information Systems Analysis and Synthesis*, 1996, pp. 24-29, [on line]. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.54.7061&rep=rep1&type=pdf

Di Nitto, E., Ghezzi, C., Metzger, A., Papazogou, M., & Pohl, K., "A journey to highly dynamic, self-adaptive service-based applications," *Automated Software Engineering*, vol 15, issue 3-4, pp. 313-341, 2008, [on line]. Available at: https://link.springer.com/article/10.1007/s10515-008-0032-x

Dou, E., & Reascos, C., "The next step in verification testing of complex systems in automation," in *Proceedings of IEEE AUTOTESTCON*, 2015, pp. 194-198, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/7356488/

El-Say, K. M., “Maximizing the encapsulation efficiency and the bioavailability of controlled-relase cetirizine microspheres using Draper-Lin small composite design,” *Drug Design, Development and Therapy*, vol 10, p. 825, 2016, [on line]. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4771436/

Encontre, V., “Testing embedded systems: do you have the GuTs for it,” *IBM*, vol 2003, issue *November*, 2003, [on line]. Available at: http://walderson.com/IBM/Practices/ScalingAgile/core.tech.common.extend_supp-ibm/guidances/whitepapers/resources/rt_test.pdf

Eigbe, P. A., Sauser, B. J., & Boardman, J., “Soft systems analysis of the unification of test and evaluation and program management: a study of a Federal Aviation Administration's strategy,” *Systems Engineering*, vol 13, no 3, pp. 298-310, 2010, [on line]. Available at: https://pdfs.semanticscholar.org/9102/438c46f62d62a10d4394a4d9fcbc1f0ebf36.pdf

Eldh, S., Brandt, J., Street, M., Hansson, H., & Punnekkat, S., “Towards fully automated test management for large complex systems,” in *Third International Conference on Software Testing, Verification and Validation*, 2010, pp. 412-420. [on line]. Available at: http://www.es.mdh.se/pdf_publications/1773.pdf

Engel, A., Bogomolni, I., Shachar, S., & Grinman, A., “Gathering historical lifecycle quality costs to support optimizing the VVT process,” in *Proceedings of INCOSE International Symposium*, 2004, vol 14, no 1, pp. 1884-1897, [on line]. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/j.2334-5837.2004.tb00621.x

Eppinger, S. D., Joglekar, N. R., & Gorschek, T., “A taxonomy for requirements engineering management in engineering design,” *Concurrent Engineering.*, vol 20, no 2, pp. 99-109, 2012.

Eppinger, S. D., Whitney, D. E., Smith, R. P., & Gebala, D. A., "A model-based method for organizing tasks in product development," *Research in Engineering Design*, vol 6, issue 1, pp. 1-13, 1994, [on line]. Available at: https://dspace.mit.edu/bitstream/handle/1721.1/2468/swp-3569-29894176.pdf?sequence=1

Estefan, J. A., "Survey of model-based systems engienering (MBSE) methodologies," *INCOSE MBSE Focus Group*, vol 25, issue 8, pp. 1-12, 2007, [on line]. Available at: http://www.omgsysml.org/MBSE_Methodology_Survey_RevB.pdf

Fiebrandt, M., & Wilson, J., 'System of systems measurement framework in evaluation of joint mission effectiveness,' in *U.S. Air Force T&E Days*, 2008, p. 1609, [on line]. Available at: https://arc.aiaa.org/doi/abs/10.2514/6.2008-1609

Flammini, A., Ferrari, P., Marioli, D., Sisnini, E., & Taroni, A., "Wired and wireless sensor networks for industrial applications," *Microelectronics Journal*, vol 40, issue 9, pp. 1322-1336, 2009, [on line]. Available at: https://www.sciencedirect.com/science/article/pii/S0026269208004618

Forsberg, K., & Mooz, H., "The relationship of system engienering to the project cycle,' *Systems Engineering*, vol 1, issue 1, pp. 57-65, 1991, [on line]. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/j.2334-5837.1991.tb01484.x

General Accounting Office (United States), *Stronger management practices are needed to improve DoD's software-intensive weapon acquisitions*, GAO-04-393, 2004, [on line]. Available at: https://www.gao.gov/new.items/d04393.pdf

Gotel, O. C., & Finkelstein, C. W., "An analysis of the requirements traceability problem," in *Proceedings of the First International Conference on Requirements Engineering*, 1994, pp. 84-101, [on line]. Available at: http://discovery.ucl.ac.uk/749/1/2.2_rtprob.pdf

Giadrosich, D.L., *Operations Research Analysis in Test and Evaluation*, 1995, Washington, D. C., USA: AIAA Education Series.

Goos, P., & Jones, B., *optimal Design of Experiments: A Case Study Approach*, New York, NY, USA; Wiley, 2011.

Graham, D., "Requirements and testing: seven missing-link myth," *IEEE Software*, vol 18, issue 5, pp. 15-17, 2002, [on line]. Available at: https://pdfs.semanticscholar.org/da4d/ac0401305e9bf9b3f2c4d264e7e6f0779f92.pdf

Gunn-Golkin, A., Martin, A., Bippert, E., Leim, S., Fann, S. & Murphy, A., "T-38C optimal landing technique determination (project talon spot)," Air Force Flight Test Center, Edwards Air Force Base, CA, USA, AFFTC-TIM-10-01, January, 2010. [on line] Available at: www.dtic.mil/dtic/tr/fulltext/u2/a536258.pdf

Gunter, B., "Sequential experimentation," *Wiley StatsRef: Statistics Reference Online*, 2014, [on line]. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat04107

Hahn, G. J., & Meeker, W. Q., "Assumptions for statistical inference," *The American statistician,* vol 47, issue 1, pp. 1-11, 1993, [online]. Available at: https://www.tandfonline.com/doi/abs/10.1080/00031305.1993.10475924

Hardin, R. H., & Sloane, N. J. A., "A new approach to the construction of optimal designs," *Journal of Statistical Planning and Inference*, vol 37, issue 3, pp. 339-369, 1993, [on line]. Available at: http://neilsloane.com/doc/Me175.pdf

Harman, M., Cortes, L. A., & Hill, R. R., "Quantifying test risk using design of experiments," *The ITEA Journal of Test and Evaluation*, vol (2016), issue 37, pp. 160-166, 2016, [on line]. Available at: https://www.afit.edu/stat/statcoe_files/Quantifying_Test_Risk_Using_Design_of_Experiments3.pdf

Heitmeyer, C., Labaw, B., & Kiskis, D., "Consistency checking of SCR-style requirements specifications," in *Proceedings of 2nd IEEE International Symposium on Requirements Engineering*, 1995, pp. 56-63, [on line]. Available at: http://www.dtic.mil/dtic/tr/fulltext/u2/a465317.pdf

Helle, P., & Schamai, W., "Towards an integrated methodology for the development and testing of complex systems," in *Proc. 5th Int. Conf. Adv. Syst. Test. Validation Lifecycle*, 2016, pp. 55-60. [on line]. Available at: https://www.researchgate.net/profile/Philipp_Helle/publication/259240893_Towards_an_Integrated_Methodology_for_the_Development_and_Testing_of_Complex_Systems/links/544672fd0cf2f14fb80f3ca0/Towards-an-Integrated-Methodology-for-the-Development-and-Testing-of-Complex-Systems.pdf

Helmerich, A., Koch, N., Mandel, L., Braun, P., Dornbusch, P., Gruler, A., & Wild, T., "Study of worldwide trends and R&D programmes in embedded systems in view of maximizing the impact of a technology platform in the area," *Final Report of the European Commission*, 2005, Brussels, Belgium.

Hesari, S., Behjati, R., & Yue, T., "Towards a systematic requirement-based test generation framework: Industrial challenges and needs," in *21st IEEE International Requirements Engineering Conference*, 2013, pp. 261-266. [on line]. Available at: https://www.computer.org/csdl/proceedings/re/2013/9999/00/06636727.pdf

Hess, J. T., & Valerdi, R., "Test and evaluation of a SoS using a prescriptive and adaptive testing framework," in *Proceedings of 5th International Conference on System of Systems Engineering (SoSE)*, 2010, pp. 1-6, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/5544070/

Hill, R., & Chambal, S. P., "A framework for improving experimental design and statistical methods for test and evaluation," in *U.S. Air Force T&E Days*, 2009, pp. 1-5.

Hodges, W., *A shorter model theory*. Cambridge university press, 1997.

Hogg, R. V., & Craig, A. T., *Introduction to Mathematical Statistics, 5th edition*, Pearson Education, 2005.

Hoke, A. T., "Economical second-order designs based on irregular fractions of the 3^n factorial," *Technometrics*, vol 16, no 3, pp. 375-384, 1974, [on line]. Available at: https://www.tandfonline.com/doi/abs/10.1080/00401706.1974.10489206

Holt, J., Perry, S., Payne, R. J., Bryans, J., Hallersteded, S., & Hansen, F. O., "A model-based approach for requirements engineering for systems of systems," *IEEE Systems Journal*, vol 9, issue 1, pp. 252-262, 2015, [on line]. Available at: https://core.ac.uk/download/pdf/30619811.pdf

Hsueh, M. C., "Large complex system test: objectives and approaches," In *Engineering of Complex Computer Systems*, 1995, pp. 405-408. [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/479366/

Hynninen, P., & Kauppinen, M., "A/B testing: a promising tool for customer value evaluation," in *Proceedings of IEEE 1st International Workshop on Requirements Engineering and Testing (RET)*, 2014, pp. 16-17, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/6908673/

International Council on Systems Engienering, *Systems Engineering Handbook*, 4th ed. San Diego, CA, USA: Wiley 2015.

Jones, A, & Deshmukh, A., "Test beds for complex systems," *Communications of the ACM*, vol 48, issue 5, pp. 45-50, 2005, [on line]. Available at: https://dl.acm.org/citation.cfm?id=1060738

Karban, R., Zamparelli, M., Bauvir, B., Koehler, B., Noethe, L., & Balestra, A., "Exploring model based engineering for large telescopes: getting started with descriptive models," *Modeling, Systems Engineering, and Project Management for Astronomy III*, vol 7017, p. 701711, 2008, [on line]. Available at: http://www.mbse.gfse.de/documents/SPIE-SysML-with-Copyright.pdf

Kalawsky, R. S., O'Brien, J., Chong, S., Wong, C., Jia, H., Pan, H., & Moore, P. R., "Bridging the gaps in a model-based system engienering workflow by encompassing hardware-in-the-loop simulation," *IEEE Systems Journal*, vol 7, issue 4, pp. 593-605, 2013, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/6409496/

Kernstine, K. H., "Inadequacies of traditional exploration methods in systems-of-systems simulations," *IEEE Systems Journal*, vol 7, issue 4, pp. 528-536, 2013, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/6509969/

Khuri, A. I, & Cornell, J. A., *Response Surfaces, Design and Analyses*, New York, NY, USA; Marcel Dekker, 1987.

Kiefer, J., "Optimum experimental designs," *Journal of the Royal Statistical Society*, series B (methodology), pp. 272-319, 1959, [on line]. Available at: https://www.jstor.org/stable/2983802?seq=1#page_scan_tab_contents

Kline, S., *Out of the Garden: Toys, TV, and Children's Culture in the Age of Marketing*, Verso. 1995

Koshal, R. S., “Application of the method of maximum likelihood to the improvement of curves fitted by the method of moments,” *Journal of the Royal Statistical Society*, vol 96, issue 2, pp. 303-313, 1933, [on line]. Available at: https://www.jstor.org/stable/2341794?seq=1#page_scan_tab_contents

Kukula, E. P., Shaw, F. R., Wallner, R., Breckenkamp, A., Kiebuzinski, G., Nadel, L., & Wolfhope, P., “Use case mobile biometric testing and evaluation: a framework to cross-reference requirements and test methods,” in *Proceedings of IEEE Conference on Technologies for Homeland Security (HST)*, 2012, pp. 67-72, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/6459827/

Lay, D. C., *Linear Algebra and Its Applications, 3rd edition,* Boston, MA, USA; Addison Wesly, 2003.

Lenth, R. V., “Quick and easy analysis of unreplicated factorials,” *Technometrics*, vol 31, no 4, pp. 469-473, 1989, [on line]. Available at: https://www.jstor.org/stable/1269997?seq=1#page_scan_tab_contents

Li, Y. L., Roberts., B, & Grenn, M. A., “Forecasting system capability using the scenario-based experimental-design test approach,” *IEEE Systems Engineering*, 2018. Advanced online publication, doi:10.1109/JSYST.2018.2843727. Available at: https://ieeexplore.ieee.org/abstract/document/8398440/

Li, Y. L, Roberts, B., & Grenn, M. A., “Model-based systems and test engineering using system mission-oriented design-of-experiment approach,” 2018. Article submitted for publication.

Liu, Z. L., Zhang, Z., & Chen, Y., "A scenario-based approach for requirements management in engienering design," *Concurrent Engineering*, vol 20, issue 2, pp. 99-109, 2012, [on line]. Available at: https://www.researchgate.net/profile/Zhinan_Zhang/publication/254082304_A_scenario-based_approach_for_requirements_management_in_engineering_design/links/0a85e5303ab7c81e5c000000/A-scenario-based-approach-for-requirements-management-in-engineering-design.pdf

Loren, J. R., "Requirement engineering: early life cycle systems engienering for air force systems and SoS," in *Proceedings of IEEE International Conference on System of Systems Engienering (SoSE)*, 2007, pp. 1-7, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/4304284/

Meidel, K. W., "United States Air Force aircraft accident investigation board report: Ellington Field, Houston, TX, 11FEB2011," USAF tech report: T-38C,65-0337, Columbus, AFB, MS, USA, [on line]. Available at: http://www.airforcemag.com/AircraftAccidentReports/Documents/2011/021111_T-38C_EllingtonFieldTX.pdf

Montgomery, D. C., *Design and Analysis of Experiments, 7th edition*, Hoboken, NJ, USA; Wiley, 2009.

Montgomery, D.C., Myers, R.H., Carter Jr., W. H., and Vining, G. G., "The hierarchy principle in designed industrial experiments," *Quality and Reliability Engineering International*, vol 21, issue 2, pp. 197-201, 2005, [on line]. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/qre.615

Morandini, M., Penserini, L, Perini, Al, & Marchetto, A., "Engineering requirements for adaptive systems," *Requirements Engienering*, vol 22, issue 1, pp 77-103, 2017, [on line]. Available at: https://link.springer.com/article/10.1007/s00766-015-0236-0

Myers, R. H., Montgomery, D. C., *Response Surface Methodology, 2nd edition*, New York, NY, USA; Wiley, 2002.

Ncube, C., "On the engineering of systems of systems: key challenges for the requirements engineering community," in *Proceedings of Workshop on Requirements Engineering for Systems, Services and Systems-of-Systems (RES4)*, 2011, pp. 70-73, [on line]. Available at: https://www.computer.org/csdl/proceedings/ress/2011/0941/00/06043923.pdf

Nelder, J. A., "The selection of terms in response surface models – how strong is a weak-heredity principle?" *The American Statistician*, vol 52, issue 4, pp. 315-318, 1998, [on line]. Available at: https://amstat.tandfonline.com/doi/abs/10.1080/00031305.1998.10480588#.W2n2uPZFxpM

Neter, J., Wasserman, W., & Kutner, M. H., *Applied Linear Statistical Models, 2nd edition*, Homewood, IL, USA; Irwin, 1985.

Nuseibeh, B., & Easterbrook. S., "Requirements engienering: a roadmap," in *Proceedings of the Conference on the Future of Software Engienering*, 2000, pp. 35-46, [on line]. Available at: https://dl.acm.org/citation.cfm?id=336523

Ogren, I., "On principles for model-based systems engineering," *Systems Engineering*, vol 3, issue 1, pp. 38-49, 2000, [on line]. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/(SICI)1520-6858(2000)3:1%3C38::AID-SYS3%3E3.0.CO;2-B

Ott, R. L., & Longnecker, M. T., *An introduction to statistical methods and data analysis*, Nelson Education, 2015.

Oyejola, B. A., & Nwanya, J. C., "Selecting the right central composite design," *International Journal of Statistics and Applications,* vol 5, issue 1, pp. 21-30, 2015, [on line]. Available at: http://article.sapub.org/10.5923.j.statistics.20150501.04.html

Parry, G. W., "The characterization of uncertainty in probabilistic risk assessments of complex systems," *Reliability Enginering & System Safety*, vol 54, issue 2-3, pp. 119-126, 1996, [on line]. Available at: https://www.sciencedirect.com/science/article/pii/S0951832096000695

Peña, M., & Valerdi, R., "Characterizing the impact of requirements volatility on systems engineering effort," *Systems Engineering*, vol 18, issue 1, pp. 59-70, 2015, [on line]. Available at: https://dspace.mit.edu/bitstream/handle/1721.1/84023/CPS_101104_Pena_Valerdi_COCOMO.pdf.pdf?sequence=1

Piaszczyk, C., "Model based systems engienering with department of defense architectural framework," *Systems Engienering*, vol 14, issue 3, pp. 305-326, 2011, [on line]. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/sys.20180

Plackett, R. L., & Burman, J. P., "The design of optimum multifactorial experiments," *Biometrika*, pp. 305-325, 1946, [on line]. Available at: https://www.jstor.org/stable/2332195?seq=1#page_scan_tab_contents

Pyster, A., Olwell, D. H., Hutchinson, N., Enck, S., Anthony Jr, J. F., & Henry, D., "Guide to the systems engienering body of knowledge (SEBoK)', vol 1.0.1. Retrieved from: http://www.sebokwiki.org/wiki/Guide_to_the_Systems_Engineering_Body_of_Knowledge_(SEBoK)

Ramesh, B., & Jarke, M., “Toward reference models for requirements traceability,” *IEEE Transactions on Software Engineering*, vol 1., pp. 58-93, 2001, [on line]. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.40.6602&rep=rep1&type=pdf

Ramos, A. L., Ferreira, J. V., & Barceló, J., “Modle-based systems engienering: an emerging approach for modern systems,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol 42, issue 1, pp. 101-111, 2012, [on line]. Available at: https://ieeexplore.ieee.org/document/5722047/

Reichert, S. H., Morr, C., & Cooney, C. M., “Effects of heat and other factors oupon foaming properties of whey protein concentrates,” *Journal of Food Science*, vol 39, issue 1, pp. 42-48, 1974, [on line]. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2621.1974.tb00983.x

Roberts, B., Mazzuchi, T., & Sarkani, S., “Engineered resilience for complex systems as a predictor for cost overruns,” *Systems Engineering*, vol 19, issue 2, pp. 111-132, 2016, [on line]. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/sys.21339

Roquemore, K. G., “Hybrid designs for quadratic response surfaces,” *Technometrics*, vol 18, no 4, pp. 419-423, 1976, [on line]. Available at: https://www.tandfonline.com/doi/abs/10.1080/00401706.1976.10489473

Shaikh, M. A., & Moore, W. H., “Statistical interpretation of absolute system performance requirements,” *IEEE Transactions on Aerospace and Electronic Systems*, vol 33, issue 2, pp. 626-631, 1997, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/588380/

Sheehan, J. H., Deitz, P. H., Bray, B. E., Harris, B. A., & Wong, A. B., "The military missions and mans framework," No. AMSAA-TR-756, Army Materiel Systems Analysis Activity, Aberdeen Proving Ground, MD, USA, 2004, [on line]. Available at: http://www.dtic.mil/dtic/tr/fulltext/u2/a428445.pdf

Souag, A., Mazo, R., Salinesi, C., & Comyn-Wattiau, I., "Reusable knowledge in security requirements engienering: a systematic mapping study," *Requirements Engineering*, vol (2016), issue 2, pp. 251-283, 2016, [on line]. Available at: https://hal.archives-ouvertes.fr/hal-01133226/file/Reuse%20in%20SRE%20Final.pdf

Streit,A., Pham, B., & Brown, R., "A spreadsheet approach to facilitate visualization of uncertainty in information," *IEEE Transactions on Visualization and Computer Graphics*, vol 14, issue 1, pp. 61-72, 2008, [on line]. Available at: http://eprints.qut.edu.au/13647/1/04359495.pdf

Topper, J. S., & Horner, N. C., "Model-based systems engineering in support of complex systems development," *Johns Hopkins APL Technical Digest*, vol 32, issue 1, pp. 419-432, 2013, [on line]. Available at: https://s3.amazonaws.com/academia.edu.documents/38634132/32_01-Topper.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1533519696&Signature=%2B3wCePV84QrHKGLdwWSkx5yURcI%3D&response-content-disposition=inline%3B%20filename%3DModel-Based_Systems_Engineering_in_Suppo.pdf

Tsai, W. T., Bia, X., Paul, R., Shao, W., & Agarwal, V., "End-to-end integration testing design," in *Proceedings of 25th Annual International Computer Software and Applications Conference (COMPSAC 2001*), 2001, pp. 166-171, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/960613/

United States Air Force, “technical manual: deficiency reporting, investigation, and resolution,” T.O. 00-35D-54, 2015, [on line]. Available at: http://www.tinker.af.mil/Portals/106/Documents/Technical%20Orders/AFD-082216-00-35D-54.pdf

Unterkalmsteiner, M., Feldt, R., & Gorschek, T., “A taxonomy for requirements engineering and software test alignment,” *ACM Transactions on Software Engineering and Methodology*, vol 23, issue 2, p. 16, 2014, [on line]. Available at: http://www.diva-portal.org/smash/get/diva2:834207/FULLTEXT01.pdf

Uusitalo, E. J., Komssi, M., Kauppinen, M., & Davis, A. M., “Linking requirements and testing in practice,” in *Proceedings of 16th IEEE International Requirements Engineering Conference*, 2008, pp. 265-270, [on line]. Available at: https://www.researchgate.net/profile/Eero_Uusitalo/publication/221222187_Linking_Requirements_and_Testing_in_Practice/links/5492b0ab0cf225673b3e070b.pdf

van Lamsweerde, A., Darimont, R., & Leiter, E., “Managing conflicts in goal-driven requirements engineering,” *IEEE Transactions on Software Engineering*, vol 24, issue 11, pp. 908-926, 1998, [on line]. Available at: https://ieeexplore.ieee.org/abstract/document/730542/

Weiss, B.A., & Schmidt, L. C., “The multi-relationship evalaution design framework: creating evaluation blueprints to assess advanced and intelligent technologies,” in *Proceedings of the 10th Performance Metrics for Intelligent Systems Workshop*, 2010, pp. 136-145, [on line]. Available at: https://dl.acm.org/citation.cfm?id=2377603

Winkler, S., & von Pilgrim, J., “A survey of traceability in requirements engineering and model-driven development,” *Software and Systems Modeling*, vol 9, issue 4, pp. 529-565, 2009, [on line]. Available at: http://www.josra.org/res/winkler.pilgrim.2009.pdf

Wu, C. J., & Hamada, M. S., *Experiments: planning, analysis, and optimization, 2nd edition*, Hoboken, NJ, USA; Hoboken, 2009.

Xia, X., Wu, J., Lui, C., & Xu, L., "A model-driven approach for evaluating system of systems," in *Proceedings of 18th International Conference on Engineering of Complex Computer Systems (ICECCS)*, 2013, pp. 56-64, [on line]. Available at: https://www.computer.org/csdl/proceedings/iceccs/2013/5007/00/5007a056-abs.html

Zeng, B., & Tan, L., "A unified framework for evaluating test criteria in model-checking -assisted test case generation," *Information Systems Frontiers*, vol 16, issue 5, pp. 823-834, 2014, [on line]. Available at: https://link.springer.com/article/10.1007/s10796-013-9424-y

Zhang, Z., Li, X., & Liu, Z. L., "A closed-loop based framework for design requirement management," in *Proceedings of ISPE CE*, 2014, pp. 444-453, [on line]. Available at: https://books.google.com/books?hl=en&lr=&id=J5zYBAAAQBAJ&oi=fnd&pg=PA444&dq=Zhang,+Z.,+Li,+X.,+%26+Liu,+Z.+L.+(2014,+September).+A+Closed-loop+Based+Framework+for+Design+Requirement+Management.+In+ISPE+CE+(pp.+444-453).&ots=WXZDll0VME&sig=rW7ffBiNUG0VzIEFiQLHs6DueuM#v=onepage&q&f=false

Zhu, D., Prichard, E. G. D., & Silverberg, L. M., "A new system development framework driven by a model-based testing approach bridged by information flow," *IEEE Systems Journal*, 2016. Advanced online publication, doi: 10.1109/JSYST.2016.231142. Available at: https://ieeexplore.ieee.org/abstract/document/7779115/

Appendix A, Chapter 3's Raw Data Sets

Table A - 1, Raw Experimental Data for Chapter 3 Example's Elements f_1 and f_2

Run #	Collected Test Results		Input Variables (coded)				
	f_1	f_2	ξ_A	ξ_B	x_1	x_2	x_3
1	***10.82***	***80.6***	-1	1	1	-1	-1
2	***8.24***	***67.9***	-1	-1	1	1	-1
3	***9.53***	***83.1***	-1	-1	1	-1	1
4	***7.59***	***38.1***	-1	1	1	1	1
5	***11.63***	***7.7***	1	-1	1	-1	-1
6	***8.39***	***74.7***	1	1	1	1	-1
7	***13.43***	***71.2***	1	1	1	-1	1
8	***7.36***	***36.8***	1	-1	1	1	1
9	***10.27***	***81.7***	-1	-1	1	-1	-1
10	***8.36***	***66.8***	-1	1	1	1	-1
11	***12.72***	***73***	-1	1	1	-1	1
12	***8.25***	***40.5***	-1	-1	1	1	1
13	***13.63***	***74.9***	1	1	1	-1	-1
14	***8.55***	***74.2***	1	-1	1	1	-1
15	***12.84***	***63.5***	1	-1	1	-1	1
16	***8.51***	***42.8***	1	1	1	1	1
17	***12.83***	***80.9***	1	-2	0	0	0
18	***6.51***	***42.4***	1	2	0	0	0
19	***12.2***	***73.4***	1	0	-2	0	0
20	***9.82***	***45***	1	0	2	0	0
21	***8.84***	***66***	1	0	0	-2	0
22	***11.5***	***71.7***	1	0	0	2	0
23	***10.81***	***77.5***	1	0	0	0	-2
24	***13.1***	***76.3***	1	0	0	0	2
25	***12.13***	***67.4***	1	0	0	0	0
26	***11***	***86.5***	1	0	0	0	0
27	***No Data***	***77.4***	1	0	0	0	0
28	***11.8***	***74.6***	1	0	0	0	0
29	***11.83***	***79.8***	1	0	0	0	0
30	***11.2***	***78.3***	1	0	0	0	0
31	***11.8***	***74.8***	1	0	0	0	0
32	***12***	***80.9***	1	0	0	0	0

Note: Half 2^{5-1} Central Composite Design per Richert, 1974.

Table A - 2, Raw Experimental Data for Chapter 3 Example's Element f_3

Run #	Collected Test Results	Input Variables (coded)		
	f_3	ξ_A	ξ_B	x_4
1	***80.6***	-1	-1	0
2	***67.9***	1	-1	0
3	***83.1***	-1	1	0
4	***38.1***	-1	1	0
5	***79.9***	-1	0	-1
6	***74.7***	1	0	-1
7	***71.2***	-1	0	1
8	***36.8***	1	0	1
9	***81.7***	0	-1	-1
10	***66.8***	0	1	-1
11	***73***	0	-1	1
12	***40.5***	0	1	1
13	***74.9***	0	0	0
14	***74.2***	0	0	0
15	***63.5***	0	0	0

Note: Full 2^3 Box Behnken Design per Myers & Montgomery, 2002.

Table A - 3, Raw Experimental Data for Chapter 3 Example's Element f_4

Run #	Collected Test Results	Input Variables (coded)			
	f_4	ξ_A	ξ_B	x_5	x_6
1	***30.3***	1	1	1	-1
2	***31.7***	-1	1	1	1
3	***34.5***	1	-1	1	1
4	***18.5***	1	1	-1	-1
5	***29.8***	-1	-1	1	-1
6	***21.3***	-1	1	-1	1
7	***21.2***	1	-1	-1	1
8	***20.3***	-1	-1	-1	-1
9	***24.8***	-1.68	0	0	0
10	***19.8***	1.68	0	0	0
11	***23.7***	0	-1.68	0	0
12	***23.9***	0	1.68	0	0
13	***18.5***	0	0	-1.68	0
14	***33.4***	0	0	1.68	0
15	***20.1***	0	0	0	-1.68
16	***27.6***	0	0	0	1.68
17	**25.3**	**0**	**0**	**0**	**0**
18	**26.1**	**0**	**0**	**0**	**0**

Note: Small Composite Design per El-Say, 2016.