

Winter 3-21-2016

# Consistency Analysis for Judgment Quantification in Hierarchical Decision Model

Mustafa Sulaiman Abbas  
*Portland State University*

Let us know how access to this document benefits you.

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/open\\_access\\_etds](https://pdxscholar.library.pdx.edu/open_access_etds)

 Part of the [Operations Research, Systems Engineering and Industrial Engineering Commons](#)

---

## Recommended Citation

Abbas, Mustafa Sulaiman, "Consistency Analysis for Judgment Quantification in Hierarchical Decision Model" (2016). *Dissertations and Theses*. Paper 2699.

[10.15760/etd.2695](https://pdxscholar.library.pdx.edu/open_access_etds/10.15760/etd.2695)

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).

Consistency Analysis for Judgment Quantification in Hierarchical Decision Model

by

Mustafa Sulaiman Abbas

A dissertation submitted in partial fulfillment of the  
requirements for the degree of

Doctor of Philosophy  
in  
Technology Management

Dissertation Committee:  
Dr. Dunder F. Kocaoglu, Chair  
Dr. Timothy R. Anderson  
Dr. Tugrul U. Daim  
Dr. Robert L. Fountain  
Dr. Kenny Phan

Portland State University  
2016

© 2016 Mustafa Sulaiman Abbas

## **ABSTRACT**

The objective of this research is to establish consistency thresholds linked to alpha ( $\alpha$ ) levels for HDM's (Hierarchical Decision Model) judgment quantification method. Measuring consistency in order to control it is a crucial and inseparable part of any AHP/HDM experiment. The researchers on the subject recommend establishing thresholds that are statistically based on hypothesis testing, and are linked to the number of decision variables and  $\alpha$  level. Such thresholds provide the means with which to evaluate the soundness and validity of an AHP/HDM decision. The linkage of thresholds to  $\alpha$  levels allows the decision makers to set an appropriate inconsistency tolerance compatible with the situation at hand. The measurements of judgments are unreliable in the absence of an inconsistency measure that includes acceptable limits. All of this is essential to the credibility of the entire decision making process and hence is extremely useful for practitioners and researchers alike. This research includes distribution fitting for the inconsistencies. It is a valuable and interesting part of the research results and adds usefulness, practicality and insight. The superb fits obtained give confidence that all the statistical inferences based on the fitted distributions accurately reflect the HDM's inconsistency measure.

## **ACKNOWLEDGEMENTS**

All gratitude and thanks are due to Almighty Allah who has endowed me with all the capacities and abilities to be where I am today. Moreover, he has made me realize my need to be thankful to Him. For that, and all His countless blessings I am grateful to Him with every fiber of my being

I would like to express my sincere gratitude to Dr. Dundar Kocaoglu, my advisor and chairman of my dissertation committee, for his invaluable and steadfast support, guidance, and mentoring. His sage advice and constant encouragement have made my PhD experience enjoyable and facilitated my achievement.

I am very thankful to Abdenacer Moussaoui of Hewlett Packard for his undying support and help throughout my PhD journey. It is always done with a cheerful and generous attitude and delivered in an intelligent and insightful fashion. He always knows how to ask the right question.

Special thanks go to Gary Oliver of Alpha Omega Computer Systems for his essential and generous assistance in writing the simulation software. His smart and ingenious application of the solutions made the research tool powerful, and enabled reliable and accurate results.

I am enormously grateful to our department chairman Dr. Tim Anderson for his invaluable and relentless efforts to correct and improve both the content and form of my dissertation. I am especially appreciative of his guidance early in my research which put me on the right track.

I am very grateful to Dr. Bob Fountain for his piercing insight, always constructive input, and generosity in sharing his vast scholarly knowledge. Beyond that, his polite and sincere manners are always appreciated.

I am grateful to all the members of my dissertation committee, Dr. Dundar Kocaoglu, Dr. Tim Anderson, Dr. Tugrul Daim, Dr. Bob Fountain, and Dr. Kenny Phan. They have graciously and generously shared their precious time.

Shawn Wall, the manager of our department's office, has always been a "pillar" of help and support for which I am very grateful.

I am very thankful to Becky Fidler, our PhD program coordinator, who is always proactive, offers help, and follows up even without being asked.

Last but not least, I am truly indebted to my wife, Taghreed. Her keen interest in my research was inspiring. Her unconditional support, countless sleepless nights and infectious optimism fueled my success.

## TABLE OF CONTENTS

<b>ABSTRACT.....</b>	<b>i</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>ii</b>
<b>LIST OF TABLES.....</b>	<b>viii</b>
<b>LIST OF FIGURES.....</b>	<b>x</b>
<b>1 INTRODUCTION.....</b>	<b>1</b>
1.1 Background .....	1
1.2 Research Objective .....	3
1.3 The Hierarchical Decision Model (HDM) .....	4
1.4 Judgment Quantification in HDM .....	6
1.4.1 Consistency/Inconsistency .....	11
1.4.2 Definition of Inconsistency Measure in HDM .....	13
1.4.3 HDM Method Detail – Constant Sum (CS) Example.....	14
1.5 Judgment Quantification in AHP.....	19
1.5.1 Saaty’s Discrete Pairwise Comparison “Absolute” Scale .....	20
1.5.2 Right Eigenvector (REV) Method Detail .....	22
1.6 Multiplicative versus Additive Pairwise Comparison Matrix (PCM) .....	25
1.7 Other Judgment Quantification Methods .....	27
<b>2 LITERATURE REVIEW .....</b>	<b>29</b>
2.1 Applicability of AHP Research to HDM .....	29
2.2 Establishing the Significance of the Research Area .....	31
2.2.1 Widespread Use of HDM and AHP .....	31
2.2.2 The Topic of Consistency in HDM/AHP .....	32



2.2.3	Use of Simulation in HDM/AHP Research .....	33
2.2.4	HDM Applications.....	36
2.2.5	AHP Applications .....	43
2.3	Statistical Simulation Studies on the Random Index (RI) .....	44
2.4	Issues Related to AHP's Consistency Measure .....	50
2.5	Other Consistency Measures .....	54
2.6	Analysis of Research Gaps .....	66
2.7	Statistical Approach to Establishing Consistency Thresholds.....	72
2.8	Summary of Literature Review .....	77
<b>3</b>	<b>RESEARCH.....</b>	<b>78</b>
3.1	Research Approach .....	78
3.2	Research Procedure .....	81
3.2.1	Sample Sizes .....	81
3.2.2	High Quality Random Number Generator .....	83
3.2.3	Reducing Computational Intensity.....	84
3.2.4	Programming Language and User Interface .....	85
3.2.5	Procedural Efficiencies in Judgment Quantification Method .....	87
3.2.6	The Need for a Leap in Efficiency .....	87
3.2.7	Producing a Single Permutation Given its Index .....	89
3.2.8	Implementing the Stratified Sampling .....	94
3.2.9	Testing Stratified Sampling .....	94
<b>4</b>	<b>RESEARCH RESULTS .....</b>	<b>96</b>
4.1	Design of Simulation Study .....	96
4.2	Interpretation of the K-S and other GOF Tests.....	97
4.3	Summary of Results .....	100

4.4 Fitted Distributions .....	108
4.4.1 Using the Johnson SB CDF and Quantile Functions .....	111
4.5 Discussion of Results.....	115
4.5.1 Note on the Distribution for $n = 3$ .....	115
4.5.2 Change of RSV's Behavior at $n = 13$ .....	116
4.6 Assessment of Current Consistency Measure .....	122
<b>5 CONCLUSION .....</b>	<b>125</b>
<b>6 CONTRIBUTION.....</b>	<b>126</b>
<b>7 LIMITATIONS .....</b>	<b>127</b>
<b>8 FUTURE WORK.....</b>	<b>130</b>
<b>REFERENCES.....</b>	<b>131</b>
<b>APPENDIX A: PDF for <math>n = 3 - 12</math>.....</b>	<b>140</b>
<b>APPENDIX B: CDF for <math>n = 3 - 12</math> .....</b>	<b>151</b>

## LIST OF TABLES

Table 1.1: HDM Matrix A – Symbolic Representation .....	8
Table 1.2: HDM Matrix B – Symbolic Representation .....	8
Table 1.3: HDM Matrix C – Symbolic Representation .....	9
Table 1.4: Pairwise Comparisons for Example in HDM’s Constant Sum Scale .....	14
Table 1.5: HDM Matrix A – CS Numerical Example .....	14
Table 1.6: HDM Matrix B – CS Numerical Example.....	14
Table 1.7: HDM Matrix C Corresponding to the “ABCD” Orientation – CS Numerical Example.....	15
Table 1.8: Combination of All Unique Columns of All of the C Matrices - CSM Numerical Example.....	15
Table 1.9: Orientation Table - CSM Numerical Example .....	17
Table 1.10: PCM Proposed by Saaty .....	19
Table 1.11: Pairwise Comparison Matrix (PCM) – Symbolic Representation.....	19
Table 1.12: Saaty’s 9-point “Absolute” Scale.....	20
Table 1.13: Saaty’s Random Index (RI) .....	21
Table 1.14: Pairwise Comparison Values in HDM CS Scale.....	22
Table 1.15: Conversion Table Used to Transform HDM CS Values to Absolute Scale Values.....	23
Table 1.16: REV Values Converted from HDM CS Values .....	23
Table 1.17: Example of PCM for REV in Fractional Format .....	23
Table 1.18: Example of PCM for REV in Decimal Format.....	23
Table 1.19: Transpose of Matrix B from HDM Numerical Example.....	23
Table 1.20: Weight Vectors for REV Numerical Example .....	24
Table 1.21: Consistency Calculations for the REV Numerical Example .....	24
Table 1.22: List of Pairwise Comparison Prioritization Procedures.....	28
Table 2.1: Summary of AHP Simulation Studies .....	35
Table 2.2: Random Index (RI) Studies Related to REV .....	49
Table 2.3: Thresholds for the Sum of Inverse Column Sums (SICS) Method.....	56
Table 2.4: List of Research Gaps .....	66
Table 3.1: Construction of Lehmer Code of a Permutation and the Permutation’s Recovery from its Lehmer Code .....	91
Table 4.1: Inconsistency Threshold Limits for 3 – 12 Decision Variables at $\alpha = 0.01$ to $\alpha = 0.25$ .....	103

Table 4.2: Inconsistency Threshold Limits for 3 – 12 Decision Variables at $\alpha = 0.26$ to $\alpha = 0.50$ .....	104
Table 4.3: Goodness of Fit for Variables 3-12.....	105
Table 4.4: Data Verification – Average Absolute Deviation in First 50 Percentiles.....	106
Table 4.5: Fitted Distribution Parameters .....	110
Table 4.6: Average Variance and Related Calculations for $n = 3 - 13$ .....	120
Table 4.7: Orientation Table from HDM Numerical Example in Section 1.4.3 .....	123
Table 7.1: List of Maximum Number of Alternatives Evaluated All Together in Each of the HDM Application Articles.....	129

## LIST OF FIGURES

Figure 1: MOGSA Structure of a Typical HDM .....	5
Figure 2: Example of Pairwise Assignment Form.....	7
Figure 3: Decimal vs. Factoradic Representation of Numbers .....	93
Figure 4: Inconsistency Threshold Limits for 3 – 12 Decision Variables.....	102
Figure 5: Average Variance for $n = 3 - 13$ .....	120
Figure 6: RSV Behavior for $n = 3 - 13$ .....	121

# 1 INTRODUCTION

## 1.1 Background

Multi Criteria Decision Analysis (MCDA) is often used to analyze intricate and complex decision problems having multiple facets. It starts with identifying criteria and alternatives related to a decision objective. Numerical measures are then used to evaluate the relative importance of alternatives with regard to the criteria. Finally, the alternatives are prioritized and ranked [1]. By using such tools, users can analyze and evaluate complex problems having conflicting priorities and at the same time make sound decisions based on rational compromise.

The Hierarchical Decision Model (HDM) [2] which is a variant of Saaty's Analytic Hierarchy Process (AHP) [3] is a widely-accepted multi-criteria decision-making tool. The first step in the application of these methods involves structuring the decision problem into levels consisting of objectives and their associated criteria. The second step involves eliciting the preferences of the decision maker (DM) through pairwise comparisons. The third step is to process the DM's input and calculate the priorities of the objectives. The final step before analyzing the decision is to check the DM's consistency. This measure ensures that the pairwise comparisons are neither random nor illogical.

For the Hierarchical Decision Model (HDM), [Cleland and Kocaoglu \[4\]](#) use a variance-based approach to calculate the inconsistency, and recommend a 10% limit above which the reliability of the expert's judgment would be considered questionable. Similarly, for the Analytic Hierarchy Process (AHP), Saaty suggests using the consistency ratio (CR) and recommends an upper limit of 10% on CR [\[3\]](#).

Saaty's fixed 10% rule has been the subject of much criticism/dispute for being too restrictive, lacking statistical justification, having no  $\alpha$  levels, and not being a function of the number of elements (decision variables) being compared.

## 1.2 Research Objective

The objective of this research is to:

1. Establish the significance of this topic in research
2. Show the research gap for HDM with regard to establishing consistency thresholds that are:
  - a. Linked to the number of variables.
  - b. Based on statistical hypothesis testing.
  - c. Linked to corresponding  $\alpha$  levels.
3. Establish how the above-mentioned research gap would be addressed.

The first 2 objectives are addressed by conducting a thorough literature review. The 3<sup>rd</sup> objective is covered by a new methodology.

The methodology used in this dissertation is based on testing the null hypothesis that the judgmental responses obtained from a respondent are random. Rejecting this null hypothesis will mean that the inconsistency of the respondent is significantly lower than what would be expected from random judgement responses.



### 1.3 The Hierarchical Decision Model (HDM)

In response to the increasing complexity of decision-making problems in a wide variety of environments, multi-attribute hierarchical decision making tools have been developed. One such method is the Hierarchical Decision Model (HDM) proposed by [Kocaoglu \[2\]](#). It is a variant of AHP.

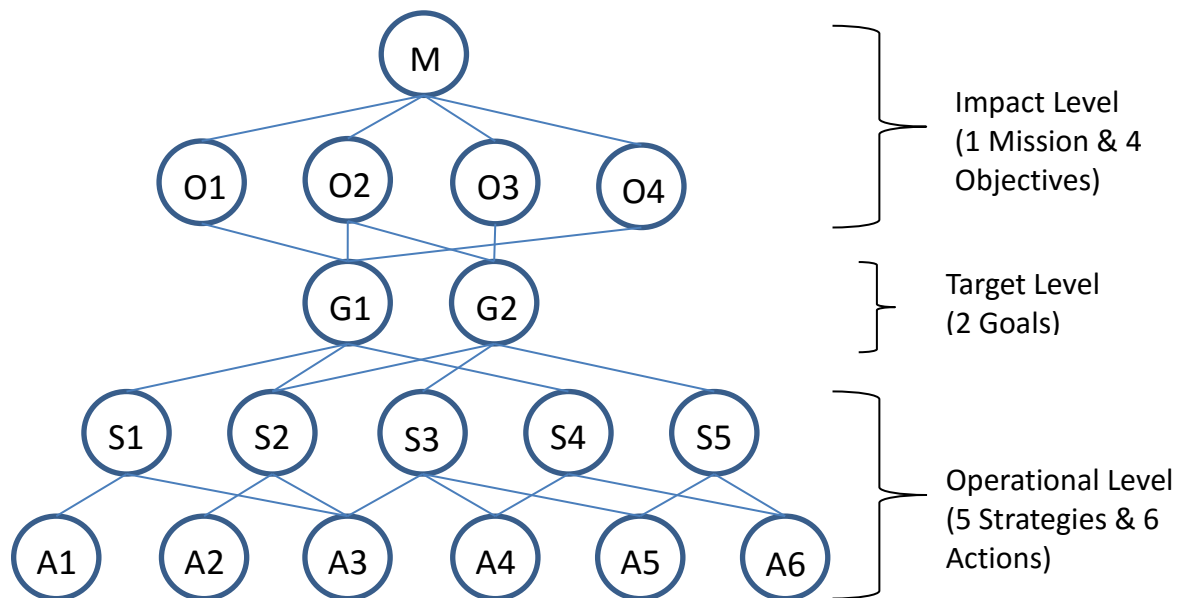
Three concepts form the basis of HDM/AHP modeling:

1. Structuring the decision problem in a hierarchy consisting of goal, criteria and alternatives.
2. Conducting pairwise comparisons among all variables at every hierarchy of the decision model with respect to each criterion on the prior/higher level.
3. Synthesis of priorities at all levels of the hierarchy after obtaining the relative judgment weights, and checking the consistency.

The first stage of building a solution in HDM is to decompose the problem into hierarchical levels at the top of which is the Mission as shown in Figure 1. The bottom level should list the alternatives under consideration. Filling the space between the top and bottom are decision criteria (attributes) that are more encompassing as they go up and less so as they go down. HDM assumes preferential independence of decision elements at each level. The construction of the hierarchical structure allows the

comparison of several alternatives on the basis of the same set of attributes. In turn, relative importance is determined.

All the second level Objectives are related to the Mission, and therefore must be compared to each other in order to determine their relative importance. However, in subsequent levels, the alternatives are not all necessarily related to all decision elements in the next higher level. Therefore, a partial selection of those alternatives could be compared to each other based on a subset of relevant elements from the next higher level.



**Figure 1: MOGSA Structure of a Typical HDM**

## 1.4 Judgment Quantification in HDM

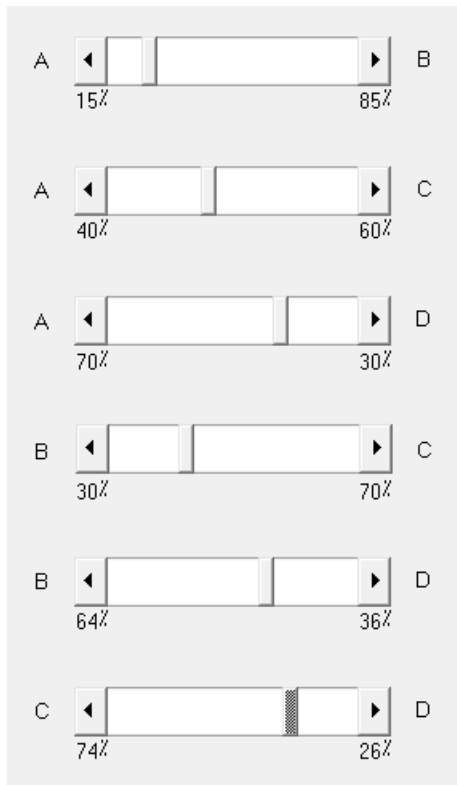
Having built the hierarchical structure, the next step is to elicit the judgment of each decision maker (DM) in a pairwise fashion. For example, if four alternatives (A, B, C, D) are considered, the pairwise comparisons would be (A : B), (A : C), (A : D), (B : C), (B : D), (C : D). HDM offers 3 input conventions (scales) for the pairwise comparisons:

1. Constant Sum (CS): The respondent is asked to divide 100 points between the two alternatives proportional to their relative values in comparison to each other with respect to the decision element under which they are being evaluated.
2. Direct Ratio (DR): The respondent is asked to provide a ratio of the two alternatives which is proportional to their respective relative values.
3. Absolute Value (AV): The respondent is asked to allocate any number of points to the two alternatives provided that the allocation is proportional to their relative values.

HDM pairwise comparisons can be given in a single format of the above or a combination of them. This allows HDM users speed and flexibility without having to mentally or arithmetically convert data into a particular scale. The fine gradations, afforded by these input scales, allow better control and accuracy without the limitation

of Saaty's nine-point scale [5]. This also prevents the negative effects of discretization often associated with Saaty's scale [6, 7].

The first step in HDM's data gathering is pairwise judgment solicitation. This is usually done through a software program that presents the DM with a list of pairwise comparisons of the criteria or alternatives which are part of the HDM model. The DM is asked to assign relative values to the alternatives. Figure 1.1 shows an example of an assignment form for pairwise comparison value judgements.



**Figure 2: Example of Pairwise Assignment Form**

The next step in HDM is priority calculations. In HDM, there are 3 matrices, simply named Matrix A, Matrix B, and Matrix C. The representations below assume that 4 elements (A, B, C, D) are being compared.

The raw input data are arranged into Matrix A in a column-oriented fashion. In Matrix A,  $C_D$ , for example, means the relative value of C when compared to D. Matrix A accepts input in any of the three formats or any combination of them.

	A	B	C	D
A	-	$B_A$	$C_A$	$D_A$
B	$A_B$	-	$C_B$	$D_B$
C	$A_C$	$B_C$	-	$D_C$
D	$A_D$	$B_D$	$C_D$	-

**Table 1.1: HDM Matrix A – Symbolic Representation**

Matrix B contains the ratios of the compared elements arranged in a column-oriented fashion as well. This is identical to Direct Ratio inputs. The elements of Matrix B above and below the main diagonal are reciprocals.

	A	B	C	D
A	1	$B_A/A_B$	$C_A/A_C$	$D_A/A_D$
B	$A_B/B_A$	1	$C_B/B_C$	$D_B/B_D$
C	$A_C/C_A$	$B_C/C_B$	1	$D_C/C_D$
D	$A_D/D_A$	$B_D/D_B$	$C_D/D_C$	1

**Table 1.2: HDM Matrix B – Symbolic Representation**

Matrix C gives the direct and indirect ratio relationships among the elements and is obtained by dividing each column of matrix B by the next column. The average of each column in Matrix C is used in the next step of calculations.

	A/B	B/C	C/D
A	$1 \div B_A/A_B$	$B_A/A_B \div C_A/A_C$	$C_A/A_C \div D_A/A_D$
B	$A_B/B_A \div 1$	$1 \div C_B/B_C$	$C_B/B_C \div D_B/B_D$
C	$A_C/C_A \div B_C/C_B$	$B_C/C_B \div 1$	$1 \div D_C/C_D$
D	$A_D/D_A \div B_D/D_B$	$B_D/D_B \div C_D/D_C$	$C_D/D_C \div 1$

**Table 1.3: HDM Matrix C – Symbolic Representation**

The next step in calculations is the construction of the orientation table. The first column of this table lists all the orientations (permutations of elements) the count of which is  $n$  factorial. The table also lists two sets of values for all elements. The first set is before normalization and the second is normalized values. Normalization means adding the values of the elements then dividing each of them by the sum. This way the normalized values of the elements will sum to unity.

To calculate the values of the elements of a single orientation, the last decision element of the orientation corresponding to Matrix C is set to 1, the values of the remaining elements are calculated based on the Matrix C ratios and their values are normalized. This is the vector of ratio scale values for the decision elements in the given orientation, (ABCD) in this example. Note that HDM's judgement quantification procedure calls for

the construction of a single Matrix C for each of the  $n$  factorial orientations involved in the calculations.

The derivation of priorities in HDM is based on the enumeration of all possible orientations of the ordering of decision elements (variables) being compared. There are  $n$  factorial orientations for  $n$  variables, such as ABCD, ACBD, ADBC, BACD, BADC..., etc. Each variable is evaluated  $n$  factorial times (once for each orientation). The weight of the variable is the arithmetic mean of these  $n$  factorial values. The normalized variable weights form the weight vector.

### **1.4.1 Consistency/Inconsistency**

Inconsistency is a slight or gross, deliberate or unintentional error in the elicited pairwise judgment related to the rank order and mutual preference proportionality of alternatives.

There are two types of consistency/inconsistency, ordinal and cardinal. Ordinal consistency requires order of preference of the ranked elements to be maintained. For example, if alternative A is preferred over B, and B is preferred over C, then A must be preferred over C. If, in this example, a user chooses C as preferred over A, then ordinal consistency is violated.

In addition to ordinal requirement, cardinal consistency requires preservation of preference proportions. For example, if A is preferred twice over B, and B is preferred thrice over C, then A must be preferred 6 times over C. If in this example, a user chooses A to be 5 times preferred over C, then cardinal consistency is violated.

It is important to note that if cardinal consistency is satisfied, then ordinal consistency, by definition, is satisfied as well, but not vice versa. Nonetheless, people in their decision-making, or when expressing their judgment are not always perfectly consistent.



Consequently, the final result will contain a certain level of inconsistency which must be measured and controlled to ensure soundness of the decision.

It is hard to overemphasize the importance of consistency in any pairwise prioritization procedure. It is a necessary parameter to ensure the reasonableness and accuracy of the prioritization result which builds confidence in both the decision and its maker.

Some but not all pairwise ranking methods provide measures for consistency/inconsistency. In fact, few judgment quantification methods offer their own consistency measures. Moreover, many consistency measures available lack meaningful interpretation because of the absence of justifiable thresholds [\[8\]](#) [\[9\]](#).

#### 1.4.2 Definition of Inconsistency Measure in HDM

In HDM, each of the  $n$  variables is evaluated  $n$  factorial times based on the ratios derived in matrix “C”. The arithmetic mean of the values is the weight of the variable. The normalized weights of the variables make up the weight vector. The variance of the mean among the values of a single variable is calculated and the sum of variance is computed. The inconsistency measure for HDM proposed in this research is the square root of the sum of variances. Therefore, the inconsistency is defined as the Root of the Sum of Variances (RSV) of the  $n$  decision elements:

$$RSV = \sqrt{\sum_{i=1}^n \sigma_i^2} \quad (1.1)$$

where  $\sigma_i^2$  is the variance of the mean of the  $i^{th}$  decision element, and  $n$  is the number of decision elements:

$$\sigma_i = \sqrt{\frac{1}{n!} \sum_{j=1}^{n!} (x_{ij} - \bar{x}_{ij})^2} \quad \forall i = 1, \dots, n \quad (1.2)$$

where  $x_{ij}$  is the normalized relative value of the variable  $i$  for the  $j^{th}$  orientation in  $n$  factorial orientations, and  $\bar{x}_{ij}$  is the mean of the normalized relative value of the variable  $i$  for the  $j^{th}$  orientation:

$$\bar{x}_{ij} = \frac{1}{n!} \sum_{j=1}^{n!} x_{ij} \quad (1.3)$$

### 1.4.3 HDM Method Detail – Constant Sum (CS) Example

The following set of pairwise comparison values is used as an example to illustrate the details of HDM’s judgment quantification method. HDM’s constant sum scale input mode is chosen for this example. The same set of pairwise comparison values will later be converted into Saaty’s REV scale and used in a similar fashion to illustrate the details of REV judgment quantification method.

<b>A</b>	30	<b>B</b>	70
<b>A</b>	47	<b>C</b>	53
<b>A</b>	65	<b>D</b>	35
<b>B</b>	79	<b>C</b>	21
<b>B</b>	63	<b>D</b>	37
<b>C</b>	17	<b>D</b>	83

**Table 1.4: Pairwise Comparisons for Example in HDM’s Constant Sum Scale**

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
<b>A</b>	-	70	53	35
<b>B</b>	30	-	21	37
<b>C</b>	47	79	-	83
<b>D</b>	65	63	17	-

**Table 1.5: HDM Matrix A – CS Numerical Example**

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
<b>A</b>	1.00	2.33	1.13	0.54
<b>B</b>	0.43	1.00	0.27	0.59
<b>C</b>	0.89	3.76	1.00	4.88
<b>D</b>	1.86	1.70	0.20	1.00

**Table 1.6: HDM Matrix B – CS Numerical Example**

	A/B	B/C	C/D
A	0.43	2.07	2.09
B	0.43	3.76	0.45
C	0.24	3.76	0.20
D	1.09	8.31	0.20
Mean	0.55	4.48	0.74

**Table 1.7: HDM Matrix C Corresponding to the “ABCD” Orientation – CS Numerical Example**

Note that HDM’s judgment quantification procedure calls for calculating a single C

Matrix for each orientation. Many of the columns of these matrices are repetitious.

Table 1.8 below combines only the unique columns of all of the C matrices for this numerical example.

	A/B	A/C	A/D	B/A	B/C	B/D	C/A	C/B	C/D	D/A	D/B	D/C
A	0.43	0.89	1.86	2.33	2.07	4.33	1.13	0.48	2.09	0.54	0.23	0.48
B	0.43	1.61	0.73	2.33	3.76	1.70	0.62	0.27	0.45	1.37	0.59	2.21
C	0.24	0.89	0.18	4.24	3.76	0.77	1.13	0.27	0.20	5.51	1.30	4.88
D	1.09	9.07	1.86	0.92	8.31	1.70	0.11	0.12	0.20	0.54	0.59	4.88
Mean	0.55	3.11	1.16	2.46	4.48	2.13	0.75	0.28	0.74	1.99	0.68	3.11

**Table 1.8: Combination of All Unique Columns of All of the C Matrices - CSM Numerical Example**

Table 1.9, shown on the next page, is the orientation table. This table lists the complete set of permutations of all the compared alternatives. These permutations are referred to as orientations. In this example, the orientations are lexicographically ordered for ease of illustration. This is not a requirement of the judgment quantification procedure. The value of an alternative is derived either directly from a single ratio or indirectly from

several ratios. The different orientations allow for the evaluation of alternatives in different combinations of direct and indirect fashions.

Before normalization							Normalized			
	Orientation	A	B	C	D	Sum	A	B	C	D
1	ABCD	1.81	3.31	0.74	1.00	6.85	0.26	0.48	0.11	0.15
2	ABDC	3.61	6.62	1.00	3.11	14.35	0.25	0.46	0.07	0.22
3	ACBD	1.88	2.13	0.60	1.00	5.61	0.34	0.38	0.11	0.18
4	ACDB	1.56	1.00	0.50	0.68	3.73	0.42	0.27	0.13	0.18
5	ADBC	3.50	4.48	1.00	3.03	12.00	0.29	0.37	0.08	0.25
6	ADCB	1.02	1.00	0.28	0.88	3.19	0.32	0.31	0.09	0.28
7	BACD	2.30	5.65	0.74	1.00	9.69	0.24	0.58	0.08	0.10
8	BADC	3.60	8.84	1.00	3.11	16.56	0.22	0.53	0.06	0.19
9	BCAD	1.16	3.86	0.86	1.00	6.88	0.17	0.56	0.13	0.15
10	BCDA	1.00	6.58	1.47	1.99	11.04	0.09	0.60	0.13	0.18
11	BDCA	1.00	4.94	0.75	2.32	9.01	0.11	0.55	0.08	0.26
12	BDAC	3.11	13.17	1.00	6.19	23.47	0.13	0.56	0.04	0.26
13	CABD	1.16	2.13	0.87	1.00	5.16	0.23	0.41	0.17	0.19
14	CADB	0.78	1.00	0.58	0.68	3.04	0.26	0.33	0.19	0.22
15	CBAD	1.16	2.84	0.81	1.00	5.80	0.20	0.49	0.14	0.17
16	CBDA	1.00	4.23	1.20	1.99	8.42	0.12	0.50	0.14	0.24
17	CDAB	0.55	1.00	0.80	1.09	3.43	0.16	0.29	0.23	0.32
18	CDBA	1.00	2.46	1.23	1.66	6.34	0.16	0.39	0.19	0.26
19	DACB	0.88	1.00	0.28	1.76	3.92	0.23	0.25	0.07	0.45
20	DABC	2.44	4.48	1.00	4.86	12.78	0.19	0.35	0.08	0.38
21	DBAC	3.11	7.65	1.00	5.17	16.93	0.18	0.45	0.06	0.31
22	DBCA	1.00	3.34	0.75	2.26	7.35	0.14	0.45	0.10	0.31
23	DCAB	0.55	1.00	0.41	1.27	3.22	0.17	0.31	0.13	0.39
24	DCBA	1.00	2.46	0.70	2.17	6.32	0.16	0.39	0.11	0.34
Mean							0.21	0.43	0.11	0.25
$\sigma^2$							0.00591	0.0106	0.00218	0.00710

**Table 1.9: Orientation Table - CSM Numerical Example**

**Example calculations for the “ABCD” orientation:**

$$D = 1$$

$$C = \frac{C}{D} \text{ (From column \#9 of table 1.8 = 0.73912)}$$

$$C = 0.74$$

$$B = C \times \frac{B}{C} \text{ (From column \#5 of table 1.8 = 4.47655)}$$

$$B = 0.74 \times 4.48 = 3.31 \text{ (3.308706)}$$

$$A = B \times \frac{A}{B} \text{ (From column \#1 of table 1.8 = 0.545894)}$$

$$A = 3.31 \times 0.55 = 1.81 \text{ (1.806202)}$$

**Inconsistency calculations:**

$$RSV = \sqrt{\sum_1^n \sigma_n^2} \quad (1.1)$$

$$RSV = \sqrt{0.00591 + 0.0106 + 0.00218 + 0.00710}$$

$$RSV = 0.161$$

## 1.5 Judgment Quantification in AHP

Saaty recommends a 9-point scale for expressing the pairwise preference and the principal or right eigenvector (REV) method for ranking the alternatives [10]. The data elicited from the decision maker (DM) is organized in the “Pairwise Comparison Matrix” (PCM).

$$\begin{array}{c} A_1 \quad A_2 \quad \cdots \quad A_n \\ \begin{array}{c} A_1 \\ A_2 \\ \vdots \\ A_n \end{array} \begin{bmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{bmatrix} \end{array}$$

**Table 1.10: PCM Proposed by Saaty**

Note that in the above PCM proposed by Saaty [10],  $w_1/w_2$  is the ratio of the weight of element  $A_1$  compared to element  $A_2$ . The main diagonal elements of the PCM are by definition 1's. This is the same as Matrix B in HDM except this is done in a row-oriented fashion.

	A	B	C	D
A	1	$A_B/B_A$	$A_C/C_A$	$A_D/D_A$
B	$B_A/A_B$	1	$B_C/C_B$	$B_D/D_B$
C	$C_A/A_C$	$C_B/B_C$	1	$C_D/D_C$
D	$D_A/A_D$	$D_B/B_D$	$D_C/C_D$	1

**Table 1.11: Pairwise Comparison Matrix (PCM) – Symbolic Representation**



### 1.5.1 Saaty's Discrete Pairwise Comparison "Absolute" Scale

**Table 1.12: Saaty's 9-point "Absolute" Scale**

Degree of Preference	Alternative A	Alternative B
According to preset interpretations of the 9 points suggested by Saaty	Any integer value in the range (1-9)	The reciprocal of A: (1/A)
<b>Example:</b> A is strongly preferred over B	Value of A = 5	Value of B = 1/5
Degree of Preference	Alternative A	Alternative B
<b>Equally preferred</b>	<b>1</b>	<b>1</b>
Equally to moderately preferred	2	1/2
<b>Moderately preferred</b>	<b>3</b>	<b>1/3</b>
Moderately to strongly preferred	4	1/4
<b>Strongly preferred</b>	<b>5</b>	<b>1/5</b>
Strongly to very strongly preferred	6	1/6
<b>Very strongly preferred</b>	<b>7</b>	<b>1/7</b>
Very strongly to extremely preferred	8	1/8
<b>Extremely preferred</b>	<b>9</b>	<b>1/9</b>

For all the choices in the above table, except equal preference (1 and 1), alternative A is preferred to alternative B. Saaty's scale is often abridged to 5 points rather than 9 with the bold values in the table above considered primary and the others referred to as intermediate. The REV method uses the normalized principal right eigenvector as the weight/rank vector.

To calculate consistency, Saaty suggests the following measure for REV

$$CR = \frac{CI}{RI} \quad (1.4)$$

$$CI = \frac{\lambda_{max} - n}{n - 1} \quad (1.5)$$

where CR is the consistency ratio, CI is the consistency index, RI is the random index,  $\lambda_{max}$ , is the maximum eigenvalue of the PCM, and  $n$  is the order of the PCM which is the same as the number of elements/variables being compared. RI is the average value of CI for randomly-generated matrices of the same order.

Saaty only accepts a matrix as consistent if  $CR < 0.1$  ( $CR < 10\%$ ). Below are two sets of suggested RI values. The first set was computed by [Forman \[11\]](#) with variable sample size ranging from 13,471 for  $n = 10$ , to 77,487 for  $n = 3$ . These RI numbers were reported and used by [Saaty \[12\]](#). The second set is from the latest RI study carried out by [Bozoki and Rapcsak \[13\]](#). This study used a much larger fixed sample of  $10^7$  for  $n = 3 - 10$ .

$n$	$\rightarrow$	3	4	5	6	7	8	9	10
RI [12]	$\rightarrow$	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49
RI [13]	$\rightarrow$	0.5242	0.8842	1.1087	1.2488	1.3408	1.4004	1.4505	1.4860

**Table 1.13: Saaty's Random Index (RI)**

### 1.5.2 Right Eigenvector (REV) Method Detail

The following set of pairwise comparisons was used in the previous example that illustrated the method detail of HDM's constant sum. The same set will be transformed into Saaty's REV scale values using the conversion table below. The transformed values will be used to illustrate REV's method detail.

The conversion is used to make the example more meaningful and allow comparison of results for illustrative purposes. There is no perfect conversion between any two pairwise comparison scales. In addition, any conversion will result in a slight amount of error due to discretization.

However, for the most accurate results, the conversion is skipped and the transpose of Matrix B is be used directly in REV. Both methods (with and without conversion) will be illustrated.

A	30	B	70
A	47	C	53
A	65	D	35
B	79	C	21
B	63	D	37
C	17	D	83

**Table 1.14: Pairwise Comparison Values in HDM CS Scale**

HDM CS Scale	50	67	75	80	83	86	88	89	90
Saaty's Absolute Scale	1	2	3	4	5	6	7	8	9

**Table 1.15: Conversion Table Used to Transform HDM CS Values to Absolute Scale Values**

A	1	B	2	A	1	C	1	A	2	D	1
B	4	C	1	B	2	D	1	C	1	D	4

**Table 1.16: REV Values Converted from HDM CS Values**

	A	B	C	D
A	1	1/2	1	2
B	2	1	4	2
C	1	1/4	1	1/4
D	1/2	1/2	4	1

**Table 1.17: Example of PCM for REV in Fractional Format**

	A	B	C	D
A	1	0.5	1	2
B	2	1	4	2
C	1	0.25	1	0.25
D	0.5	0.5	4	1

**Table 1.18: Example of PCM for REV in Decimal Format**

	A	B	C	D
A	1	0.43	0.89	1.86
B	2.33	1	3.76	1.70
C	1.13	0.27	1	0.20
D	0.54	0.59	4.88	1

**Table 1.19: Transpose of Matrix B from HDM Numerical Example**

For REV, the weight vector is the normalized right eigenvector:

	Using Table 1.18 (With Scale Conversion)					Using Table 1.19 (Without Scale Conversion)			
	A	B	C	D		A	B	C	D
Element									
Weight	0.23 2	0.41 2	0.11 6	0.23 2		0.219	0.400	0.117	0.264
Rank	2	1	3	2		3	1	4	2

**Table 1.20: Weight Vectors for REV Numerical Example**

For consistency, Saaty's suggested measure is:

$$CR = \frac{CI}{RI} \quad CI = \frac{\lambda_{max} - n}{n-1} \quad (1.2 \text{ and } 1.3)$$

	With Scale Conversion	Without Scale Conversion
$n$	4	4
$\lambda_{max}$	4.3860	4.4988
RI	0.8842	0.8842
CI	0.1287	0.16627
CR	14.55%	18.80%

**Table 1.21: Consistency Calculations for the REV Numerical Example**

According to both methods, CR is above the 10% upper threshold recommended by Saaty. Therefore, the pairwise comparison values would be considered too inconsistent to be used, and it is recommended for the DM to revise the pairwise assignments.

## 1.6 Multiplicative versus Additive Pairwise Comparison Matrix (PCM)

The approach used in HDM, REV and the majority of popular prioritization procedures in terms of developing and analyzing the PCM is “multiplicative”. The “additive” approach is an alternative. [Cavallo and D'Apuzzo \[14\]](#) explain as follows:

“Multiplicative PCM:  $a_{ij}$  represents the preference ratio of  $x_i$  over  $x_j$ :  $a_{ij} > 1$  implies that  $x_i$  is strictly preferred to  $x_j$ , whereas  $a_{ij} < 1$  expresses the opposite preference, and  $a_{ij} = 1$  means that  $x_i$  and  $x_j$  are indifferent.

The condition of multiplicative reciprocity is:

$$a_{ji} = \frac{1}{a_{ij}} \quad \forall i, j = 1, \dots, n \quad (1.6)$$

The condition of multiplicative consistency is:

$$a_{ik} = a_{ij}a_{jk} \quad \forall i, j, k = 1, \dots, n \quad (1.7)$$

Additive PCM:  $a_{ij}$  represents the difference of preference between of  $x_i$  and  $x_j$ :  $a_{ij} > 0$  implies that  $x_i$  is strictly preferred to  $x_j$ , whereas  $a_{ij} < 0$  expresses the opposite preference, and  $a_{ij} = 0$  means that  $x_i$  and  $x_j$  are indifferent.

The condition of additive reciprocity is:

$$a_{ji} = -a_{ij} \quad \forall i, j = 1, \dots, n \text{ \& } i \neq j \quad (1.8)$$

The condition of additive consistency is:

$$a_{ik} = a_{ij} + a_{jk} \quad \forall i, j, k = 1, \dots, n \text{ \& } i \neq j, j \neq k, i \neq k" \quad (1.9)$$

## 1.7 Other Judgment Quantification Methods

[Choo and Wedley \[15\]](#) compared the effectiveness of 18 methods for deriving preference values from pairwise comparison matrices (PCM). In the methods studied, including REV, a few presume interval-scaled values, while the majority of them use ratio-scaled values.

[Choo and Wedley \[15\]](#) used two criteria for measuring the effectiveness of the judgment quantification methods. The first is “correctness in error free cases”. This means the ability of the method to calculate -in a simple way, from column values- the correct rank order given a pairwise comparison matrix that is perfectly consistent. The second criterion is “distance minimization”. This means the ability of the method to calculate - in a more complicated fashion- the variables’ rank order while minimizing the difference among the ratios of the variables in the final rank vector compared to their ratios in the pairwise comparison matrix (PCM).

Later on, [Lin \[16\]](#) revised the work of Choo and Wedley and concluded that 3 pairs of the distance minimization methods were mathematically equivalent, and in effect 15 methods were truly unique. Preference weighted least worst square (PWLWS) is equivalent to Preference weighted least absolute error (PWLAE), Least worst square



(LWS) is equivalent to Least worst absolute error (LWAE), and Logarithmic least worst square (LLWS) is equivalent to Logarithmic least worst absolute error.

The following table lists the methods studied by [Choo and Wedley \[15\]](#).

LS	Least square
LWS	Least worst square
PWLS	Preference weighted least square
PWLWS	Preference weighted least worst square
LAE	Least absolute error consider only
LWAE	Least worst absolute error
PWLAE	Preference weighted least absolute error
PWLWAE	Preference weighted least worst absolute error
SGM / LLS	Simple geometric mean / Logarithmic least square
LLWS	Logarithmic least worst square
LLAE	Logarithmic least absolute error
LLWAE	Logarithmic least worst absolute error
SCS	Simple column sum
SNCS	Simple normalized column sum
REV	Right eigenvector
NREV	Normalized right eigenvector
LEV	Left eigenvector
PWGM	Preference weighted geometric mean

**Table 1.22: List of Pairwise Comparison Prioritization Procedures**

## **2 LITERATURE REVIEW**

### **2.1 Applicability of AHP Research to HDM**

HDM and AHP share an almost identical approach and structure. However, the two models use different mathematical methods for the calculation and aggregation of priorities, as well as the calculation of the decision maker's consistency.

The widespread popularity of AHP has sparked an intense research activity resulting in a rich research field. The subject, focus, findings and recommendations of a great deal of these studies are directly applicable and perfectly relevant to judgment quantification in HDM in general and HDM consistency treatment in particular. Following are specific justifications for the preceding statement and the consequent use of AHP research in fulfilling some of the objectives of this work.

1. Any AHP application can be directly applied using HDM. The results, as shown in the examples given in sections 1.4.3 and 1.5.2, are almost identical with the added benefits of better accuracy, speed, and flexibility thanks to HDM's scale options.
2. HDM's judgment quantification method falls in the same category as numerous prioritization procedures proposed as alternatives to Saaty's REV. Eighteen of these procedures are listed in table 1.22. All the evaluation criteria for AHP

prioritization procedures are directly applicable to HDM's prioritization procedure.

3. HDM's consistency measure falls in the same category as numerous method-specific consistency measures proposed as alternatives to Saaty's consistency measure (CR). These consistency measures are described at length in section 2.5 of this chapter. All the evaluation criteria for AHP consistency measures are directly applicable to HDM's consistency measure.
4. All of the AHP consistency research is also directly applicable to HDM's consistency approach. More specifically:
  - a. AHP consistency research provides an historical comprehensive treatment of the major topic of pairwise consistency. As a whole, the research is not limited to a particular method, a specific subtopic, or a single approach. As will be shown, numerous research papers are extremely relevant to HDM's consistency measure.
  - b. HDM's and Saaty's AHP consistency measures recommend a 10% fixed threshold. Therefore, all of the research investigating the validity of this approach applies equally well in both cases.
  - c. The various statistical analysis and simulation studies concerning AHP pairwise consistency are especially relevant to HDM's consistency approach. The ubiquitous conclusions of a great deal of such studies specifying the drawbacks of the fixed threshold approach and their

recommendations of a hypothesis testing approach as the most valid and reasonable approach form the basis of this research

## **2.2 Establishing the Significance of the Research Area**

### **2.2.1 Widespread Use of HDM and AHP**

There is widespread evidence that AHP and its variants such as HDM are some of the most important research areas in the field of decision making. Merely a decade after its proposal, even an AHP critic admitted that AHP has established itself as a “major tool in multi-criteria decision analysis” [6]. The widespread acceptance of AHP in the US and worldwide is often attributed to the power and simplicity of AHP [17, 18]. The applicability and flexibility of AHP has also contributed to its great popularity and has helped make it one of the most widely-used decision-making tools [18-21]. AHP and its variants have been applied in a multitude of fields across all sectors where decision-making is needed [18, 21-24]. All of this has given AHP “an impressive record of success” [25]. It is of great importance and relevance to point out that the popularity and success of AHP has also made it a heavily researched area in decision making [18, 19, 21, 26]. The sheer volume of research articles on AHP and its variants [21] and numerous literature reviews on the same subject [21, 26-28] clearly establish this as one of the most important areas of research in decision making science.

### **2.2.2 The Topic of Consistency in HDM/AHP**

Before the synthesis of single-level priorities or aggregate priorities across multiple levels, HDM/AHP models require assignment of preference to the various elements being compared. This is done by eliciting the input of a decision maker or an expert in pairwise comparison of the elements. Inconsistency in the choices made by a decision maker is contradiction in terms of order of preference (ordinal inconsistency) or in terms of relative degree of preference (cardinal inconsistency). Compliance with cardinal consistency leads necessarily to compliance with ordinal consistency but not vice versa. Since pairwise consistency/inconsistency can directly affect the quality and integrity of the order and degree of preference in the final result, there is consensus among decision scientists that inconsistency should be measured and controlled within an upper limit.

The importance of consistency in AHP is well stated by AHP's original author, Thomas Saaty, "how to measure inconsistency and improve the judgments to obtain better consistency is a concern of the AHP" [29]. Because the soundness of the result of an AHP model, or any pairwise comparison for that matter, is directly related to consistency, the analysis of this parameter is a critical step [30, 31], and an important consideration in AHP [32, 33]. In AHP, improving consistency improves the validity of

judgments [1]. All of these considerations have made the topic of consistency one of the most researched topics in AHP [34-38].

### **2.2.3 Use of Simulation in HDM/AHP Research**

Simulation has been extensively used in AHP research. The following examples do not include simulation use for threshold calculation. These will be discussed in a later section.

Budescu, *et al.* [39] used simulation to compare the performance of 2 prioritization procedures. Zahedi [40] used simulation to assess the performance of 6 prioritization procedures. Noble and Sanchez [41] introduced the parameter of “entropy” to measure the meaningful information contained in an AHP decision maker’s input and used simulation to show that “entropy” is normally distributed. Saaty and Vargas [42] used simulation to examine rank reversal under the 3 AHP modes: distributive, ideal, and utility. Genest and Rivest [43] used simulation to evaluate the performance of REV versus the geometric mean method (GMM) prioritization procedures. To facilitate pairwise data collection, Carmone Jr, *et al.* [44] suggest requiring a reduced set of inputs from decision makers. They used simulation to show that AHP models using their suggestion do not suffer from inaccuracy due to data loss. Finan and Hurley [45] used simulation to prove that further reduction of inconsistency in matrices already considered consistent is still beneficial to the overall decision analysis in AHP.

Zanakis, *et al.* [46] used simulation to evaluate the performance of 4 multi-attribute decision making tools in addition to 4 AHP versions. Choo and Wedley [15] used simulation to assess the performance of 18 prioritization procedures. Aull-Hyde, *et al.* [47] used simulation to prove that when using the geometric mean method for aggregation of individual judgments in AHP, the group consistency is unaffected by the inclusion of a few inconsistent judgments if the group size is sufficiently large. Ishizaka and Lusti [48] used simulation to assess the performance of 4 prioritization procedures. Lin [16] used simulation to revise the work of Choo and Wedley [15] and concluded that 3 pairs of the prioritization methods were mathematically equivalent, and in effect 15 methods were truly unique. Mamat and Daniel [49] used simulation to demonstrate the benefit of using the singular value decomposition (SVD) method in AHP, in terms of speed and reduction of the number of pairwise comparisons required. Dong, *et al.* [50] used simulation to demonstrate benefit of two proposed consensus models for AHP group decisions. Ishizaka, *et al.* [51] used simulation to demonstrate the effect of scale and aggregation on ranking of alternatives in AHP. Siraj, *et al.* [52] used simulation to show that a large proportion of PCMs that are deemed consistent according to Saaty's CR measure could actually be ordinally inconsistent. Their simulation also shows that cardinal inconsistency tends to decrease under aggregation whereas ordinal inconsistency does not.

**Table 2.1: Summary of AHP Simulation Studies**

<b>Comparison of judgment quantification methods, decision-making tools, or judgement scales</b>	
<b>Article</b>	<b>Subject</b>
<a href="#">Budescu, et al. [39]</a>	Comparison of right eigenvalue to geometric mean
<a href="#">Zahedi [53]</a>	Comparison of right eigenvalue, mean transformation, row geometric mean, column geometric mean, harmonic mean and simple row average.
<a href="#">Genest and Rivest [43]</a>	Comparison of right eigenvalue to row geometric mean
<a href="#">Choo and Wedley [15]</a>	Comparison of 18 judgment quantification methods (detailed in section 1.7)
<a href="#">Lin [16]</a>	Comparison of distance minimization methods included in <a href="#">Choo and Wedley [15]</a>
<a href="#">Ishizaka and Lusti [48]</a>	Comparison of right eigenvector, left eigenvector, geometric mean and mean of normalized values.
<a href="#">Mamat and Daniel [49]</a>	Comparison of singular value decomposition to the duality approach in AHP
<a href="#">Ishizaka, et al. [51]</a>	Comparison of the following scales: absolute, power, geometric, logarithmic, root square, inverse linear, and balanced.
<a href="#">Zanakis, et al. [46]</a>	Comparison of four multi-attribute decision making: ELECTRE, TOPSIS, Multiplicative Exponential Weighting (MEW), Simple Additive Weighting (SAW). Comparison of two AHP scales: absolute versus geometric. Comparison of two AHP methods: right eigenvector versus mean transformation.
(Continued on the next page)	



Table 2.1: Summary of AHP Simulation Studies (Continued from the previous page)	
Examination of a topic or a phenomenon related to AHP	
Article	Subject
<a href="#">Noble and Sanchez [41]</a>	Assessment of information content of pairwise comparison data using entropy
<a href="#">Saaty and Vargas [42]</a>	Rank reversal in AHP
<a href="#">Carmone Jr, et al. [44]</a>	Streamlining pairwise elicitation procedure by allowing incomplete input matrices
<a href="#">Finan and Hurley [45]</a>	Usefulness of further reduction of inconsistency of already consistent matrices
<a href="#">Aull-Hyde, et al. [47]</a>	Effect of using geometric mean on aggregate consistency
<a href="#">Dong, et al. [54]</a>	Test two proposed consensus models for AHP group decisions
<a href="#">Siraj, et al. [52]</a>	CR could admit ordinally inconsistent matrices. Cardinal inconsistency improves by aggregation whereas ordinal does not

#### 2.2.4 HDM Applications

Following are summaries of selected HDM-specific applications.

An early and interesting application of HDM to capital budgeting was carried out by [Khan \[55\]](#) who found HDM easy to use and flexible. The author related HDM attractiveness to its ability to rank discrete alternatives and their attributes without those being restricted to any particular class. Through a capital rationing example, Khan presented a valuable and thorough explanation of HDM's structure as well as its judgment quantification method for priority ranking and consistency calculations.

[Gersri \[56\]](#) combined a 4-level HDM structure with the Delphi method to develop a Technology Development Envelope (TDE) for emerging technologies. The author validated compliance of the structure with HDM's assumption of criteria independence. The author recommends this method as a systematic approach for developing TDE.

A valuable contribution to HDM's usefulness, versatility, and robustness was the development of a sensitivity analysis (SA) method by [Chen and Kocaoglu \[57\]](#). The SA algorithm provides means for calculating the effect of changes at any local level of the HDM structure on the overall rank of priorities. This expands the understanding of the relationships among the alternatives and may lead to simplification of some model complexities. The method is actually independent of the judgment quantification method employed and can be used in AHP as well as any of its variants that use additive function for priority aggregation. [Chen, et al. \[58\]](#) apply the SA developed earlier to predict future changes to industry-wide economic conditions and provide organizational strategies for dealing with them.

[Gersri and Kocaoglu \[59\]](#) use HDM to develop a systematic approach for planning R&D strategies and policies. The HDM model used integrated multiple decision levels and methodologies to develop a decision making tool suitable for effective allocation of national resources to support emerging technologies.

In order to operationalize technology roadmapping, [Fenwick, et al. \[60\]](#), use HDM to link markets to products, products to technologies, technologies to R&D projects, which are in turn linked to technology roadmapping . The structured HDM is considered over time using time-sensitive variables to finally derive the Technology Development Envelope (TDE).

[Kodali, et al. \[61\]](#) emphasize the originality of using HDM to demonstrate the shortcomings of total productive maintenance (TPM) and to show that world-class maintenance systems (WMS) is the best among the proposed “best practices” solutions. HDM helps achieve these conclusions by allowing the authors to link and analyze performance measures of an organization to its maintenance systems. In the course of using HDM, the paper provides a thorough explanation of HDM’s structure, judgment quantification method for calculating priorities, priority aggregation, and pairwise judgment consistency. The authors recommend HDM for all practitioners in the field of maintenance management.

[Cowan, et al. \[62\]](#) use a 4-level HDM to explore the impact of technology development and adoption on the sustainability objectives of hydroelectric generation & storage technologies (HPSTs) in the US Pacific Northwest. The priorities of the technology selection from the HDM structure is fed into a linear programming (LP) model to further analyze the sustainability factors in the region.

[Daim, et al. \[63\]](#) use a hybrid model of HDM and goal programming (GP) to reverse generate the energy policies of France and Germany who represent the 2 extremes of national energy mix in Europe. The interesting approach starts with setting up an HDM model which is analyzed/verified in a backward fashion by the GP model. Parts of the findings did not match those countries' energy portfolios suggesting the need for further research as well as further refinement of the model.

[Harell and Daim \[64\]](#) develop a selection tool of suitable employee motivational strategy using HDM. Instead of having the HDM model lead to a single selection of motivational strategy goal as is the case in previous research, this paper offers a motivational strategy pathway giving managers deeper insight on the design and implementation of their choices. In addition a survey of 50 professionals if carried out to demonstrate the acceptance of the tool and to provide further details regarding employee motivational priorities for groups versus subgroups.

[Kennedy and Daim \[65\]](#) use HDM to incorporate goals and aspirations of company employees into those of the company's stakeholders. Rather than wasting results from employee satisfaction surveys, which happens too often, the HDM model is intended to utilize such valuable data to enhance employee engagement and retention. The authors suggest pairwise comparison for conducting the employee surveys as a much better

alternative to the commonly used Likert scale. The advantages of doing so include the reduction of a great deal of data into manageable size as well as greater detail and resolution from the survey responses.

[Wang, et al. \[66\]](#) develop an HDM to evaluate the most suitable energy resource for China. Further sensitivity analysis (SA) is used to gauge the robustness of the model and the change in rank priorities under foreseeable circumstances. Because “current energy infrastructure” is the most critical selection criterion, coal becomes the most favored energy source. However, SA shows that as the criterion of “environmental impacts” gains priority, renewable resources will quickly surpass coal as the favored alternative.

A simple yet effective and useful HDM model for hybrid car selection is presented by [Fenwick and Daim \[67\]](#). The problem faced by many consumers lends itself to HDM structure. The model allows users to determine their priorities through 3 upper level criteria and 4 lower level ones. The model contains a database of vehicle characteristics related to model's attributes. Combining the user preferences along with available vehicle data, the model is able to provide the user with a suggested matching option. Based on car salesmen interviews, the authors suggest that future improvements of the model could include expanding the criteria numbers to include some secondary options.

To carry out assessment and selection of renewable energy generation technologies using multiple perspectives, [Sheikh, et al. \[68\]](#) use HDM as the major decision making tool. The authors compare HDM/AHP to Multi-attribute Utility Theory (MAUT), and Outranking as multi-criteria decision analysis (MCDA) tools. They state, based on several literature reviews, that HDM/AHP is the most popular tool used in the field of energy planning. The authors also conclude that HDM is most suitable for the objectives of their research because of its flexibility and scalability with regard to accommodating multiple perspectives. They also point out HDM's advantage in integrating individual and group rankings.

Building data centers (DCs) is a costly investment that many businesses must make. Therefore, [Daim, et al. \[69\]](#) propose an HDM for site selection for DCs. The main selection criteria include geographical, financial, political, and social factors. The authors gather the pairwise comparison data from experts and use the PCM software developed at Portland State University (PSU) to automate HDM's judgment quantification method. The authors recommend using the model as a valuable decision making tool that is readily available.

In order to assess the performance and ensure compliance of information systems (IS), the Korean government has established standard IS audit checks and authorized

licensed companies to carry out such audits. [Lee, et al. \[70\]](#) provide the first HDM model to evaluate the relative importance of the standard check items in an IS audit. In order to improve the quality of the audits, the author suggests using the research results as basis to introduce logical and systematic changes or modifications to the priorities of check items to suit the particular conditions of the system of interest.

### **2.2.5 AHP Applications**

This section is included to show the wide-spread use and acceptance of AHP in the field of decision making.

[Vaidya and Kumar \[27\]](#) provide this recent and comprehensive survey paper on AHP applications. In total, 150 AHP application papers were reviewed 27 of which were analyzed in depth. In addition to providing an excellent summary of the AHP application literature, this paper proves the wide acceptance of AHP as a decision making tool of choice in a multitude of truly diverse areas of application.

The usefulness of this paper is in its potential use as a guide to previous application work that may help both researchers and practitioners decide the proper fit of AHP to their own work. The chronological organization of research is reflected in the authors' selection of papers for their study. This allows researchers to track the development of concepts in the process of choosing an application method suitable for their situation.

The authors classify the reviewed papers based on their theme of application. These themes are selection, evaluation, benefit–cost analysis, allocations, planning and development, priority and ranking, decision-making, forecasting, medicine, and AHP application with QFD (Quality Function Deployment). The areas of application such as:



personal, social, manufacturing, political, engineering, education, industry, and government are combined with the theme classification to add more resolution and relevance to the classification.

The paper provides a useful chart detailing the year of publication of the articles covered. The chart shows that 12% were published prior to 1990, 27% in the period (1998 – 2000) and 31% during (2000 – 2003). Region-wise sorting of AHP applications shows USA's share as 47%, followed by Asia at 33%, then Europe at 18%. Finally, the paper shows the distribution of the reviewed papers among journals. The lion share goes to the *European Journal of Operational Research*, which incidentally is the publisher of the [Vaidya and Kumar \[27\]](#) literature review.

### 2.3 Statistical Simulation Studies on the Random Index (RI)

As mentioned previously, Saaty recommends the following consistency measure for REV

$$CR = \frac{CI}{RI} \quad CI = \frac{\lambda_{max} - n}{n-1} \quad (\text{Equations 1.2 and 1.3})$$

where CR is the consistency ratio, CI is the consistency index which is a mathematical quantity calculated from the PCM, and RI is the random index which is the average value of CI for a sample of randomly-generated matrices of the same order (number of variables). Saaty's upper limit for CR is 10%. Saaty's calculation of RI involves the

generation of randomly populated matrices, calculating their CIs, and then averaging the results of the sample. Obviously, this is done repeatedly for a range of alternatives, typically  $n = 3 - 10$ .

The initial RI numbers reported by Saaty were composite results of 2 simulation runs. The first was performed by V.R. Uppuluri (at Oak Ridge) using a sample of 100 matrices, and the second was performed by Saaty (at Wharton) using a samples of 500 matrices [71].

Lane and Verdini [72] conducted their study to check the validity of Saaty's 10% rule and examine the random distribution of the CI, REV's inconsistency measure. For  $n = 3$ , the authors generated the complete probability distribution of CI using the full enumeration of 4,913 matrices. 2500 matrices each were used for  $n = 4 - 10, 12, 14, 16, 20$ , and 24. The authors conclude that for  $n = 3$  and 4, the 10% rule seems too lax and needed to be stricter. For  $n = 3$ , they recommend an RI which corresponds to  $\alpha = 0.05$ , and for  $n = 4$ , they recommend an RI which corresponds to  $\alpha = 0.01$ . For  $n$  higher than 4, they conclude the 10% rule is much stricter than statistical rules ( $\alpha$  levels) and choose to support it. The authors justify their choice by stating that even "semi-rational" DMs are able to comply with strict  $\alpha$  levels, particularly when  $n$  is large, and therefore stricter levels were necessary to ensure quality decisions. This view is not shared by any of the other researchers.

Golden and Wang [71] conducted their study to answer, among other things, the questions: why should the threshold be set at 10%?, and should the threshold be a function of matrix size? Unlike Saaty, Uppulari, and Lane & Verdini, the authors did not randomly populate their sample matrices. Instead, they used a complex algorithm to fill their 1,000-matrix samples. The underlying assumption of the fill procedure was that DMs will always try, and mostly succeed, in being consistent. The algorithm employs a variable ( $k = 1 - 5$ ) which reflects the DM's ability to be consistent. The numbers reported in the study were based on  $k = 3$ , indicating the DM is earnestly trying to be consistent.

The RI study carried out by Forman [11] was done for incomplete matrices of dimension  $n = 3 - 7$  with  $M = 1 - 15$  missing elements. He used Harker's [73] algorithm for calculating the priorities as well as CI.

Dodd, *et al.* [74] carried out a simulation study using a sample of 1,000 matrices each for  $n = 4 - 9$ . The authors show that even at 20% of the random mean, few to none of the higher order matrices ( $n = 6 - 9$ ) passed Saaty's CR limit. Since they strongly argue against the RI/CR approach, the authors, unlike those of other RI studies, did not report the random mean of their study samples (10% of which would be considered RI). Instead, as they strongly advocate a statistical hypothesis testing approach, they

reported the permitted CI thresholds that correspond to  $\alpha$  levels of 0.1%, 0.5%, and 1%-5%.

[Tummala and Wan \[75\]](#) developed a closed form expression for  $\lambda_{max}$  for PCMs where  $n = 3$ , and used the formula to determine the mean and variance of  $\lambda_{max}$ . For higher order PCMs, the mean and variance of  $\lambda_{max}$  were determined by simulation. The authors then used these  $\lambda_{max}$  statistical parameters to generate a new set of RI numbers. Their formula for a 3 by 3 PCM:

$$\begin{pmatrix} 1 & a & b \\ 1/a & 1 & c \\ 1/b & 1/c & 1 \end{pmatrix}, \quad \lambda_{max} = 1 + \sqrt[3]{\frac{b}{ac}} + \sqrt[3]{\frac{ac}{b}} \quad (2.1)$$

In the course of developing thresholds for the consistency measure of the Geometric Mean Method, [Aguaron and Moreno-Jimenez \[8\]](#), performed an RI simulation using a sample of 100,000 matrices.

[Alonso and Lamata \[9\]](#) did their RI study in the course of developing a new consistency measure. The authors used two sets of sample sizes; the first was 100,000 matrices and the second was 500,000 matrices. They found no difference between the results of the two simulation runs.

[Bozoki and Rapcsak \[13\]](#) performed the most recent and thorough RI study. The objective of their research was to compare REV's consistency measure with the one proposed by [Koczkodaj \[76\]](#). The sample they used was  $10^7$  matrices for  $n = 3 - 10$ . The  $\lambda_{max}$  plots show it becomes close to being normally distributed particularly for higher order matrices ( $n = 6 - 10$ ). The plots also show Saaty's 10% rule well outside to the left of the distribution for the same range. For REV's CR, the authors see the 10% rule's lack of linkage to matrix size, and its inability to exclude asymmetric inconsistency as major weaknesses. For Koczkodaj's CM, the question of extending the recommended threshold to higher orders remained an open question.

It is interesting to note that despite being done 21 years earlier than the latest RI study, and despite the much smaller sample size (1,000), the results obtained by [Golden and Wang \[71\]](#) for  $n$  higher than 4 differ on average by less than 0.5% from those obtained by the 3 most recent RI studies.

Table 2.2 shown on the next page summarizes the RI studies to date. Primary source for table data is [Alonso and Lamata \[9\]](#). Sample sizes for the studies of Forman, Tumala & Wan, as well as the column (study) of Bozoki & Rapcsák were added by the author of this dissertation.

Author →	Uppuluri	Saaty	Lane & Verdini	Golden & Wang	Forman	Noble	Tumala & Wan	Aguaron, <i>et al.</i>	Alonso & Lamata	Bozóki & Rapcsák
Sample $n$ ↓	100	500	2,500	1,000	17,672 to 77,487	5,000	4,600 to 470,000	100,000	100,000	$10^7$
3	0.382	0.58	0.52	0.5799	0.5233	0.49	0.500	0.525	0.5245	0.5242
4	0.946	0.90	0.87	0.8921	0.8860	0.82	0.834	0.882	0.8815	0.8842
5	1.220	1.12	1.10	1.1159	1.1098	1.03	1.046	1.115	1.1086	1.1087
6	1.032	1.24	1.25	1.2358	1.2539	1.16	1.178	1.252	1.2479	1.2488
7	1.468	1.32	1.34	1.3322	1.3451	1.25	1.267	1.341	1.3417	1.3408
8	1.402	1.41	1.40	1.3952		1.31	1.326	1.404	1.4056	1.4004
9	1.350	1.45	1.45	1.4537		1.36	1.369	1.452	1.4499	1.4505
10	1.464	1.49	1.49	1.4882		1.39	1.406	1.484	1.4854	1.4860
11	1.576	1.51		1.5117		1.42	1.433	1.513	1.5141	
12	1.476		1.54	1.5356		1.44	1.456	1.535	1.5365	
13	1.564			1.5571		1.46	1.474	1.555	1.5551	
14	1.568		1.57	1.5714		1.48	1.491	1.570	1.5713	
15	1.586			1.5831		1.49	1.501	1.583	1.5838	

**Table 2.2: Random Index (RI) Studies Related to REV**

## 2.4 Issues Related to AHP's Consistency Measure

[Barzilai, et al. \[77\]](#) contend that using the principle right eigenvalue to calculate consistency is not justified. They claim the major fault with this approach is its dependence on the problem description. They also indicate that solving for each level separately and then combining the solutions yields different results from combining the levels and then calculating the solution. They call this discrepancy “inter-level inconsistency”. The authors develop a set of desired properties that a prioritization procedure should have: 1) If a matrix is consistent, its weight vector should be the solution, 2) The solution should be independent of the problem description, and 3) The solution should provide inter-level consistent decisions. The authors show that the Geometric Mean Method developed by [Crawford and Williams \[78\]](#) is the only method that would satisfy these properties.

[Golden and Wang \[71\]](#) were first to point out that under the recommended 10% rule for CR, some obviously inconsistent matrices could be deemed consistent. The example the authors provide is a 3x3 matrix which violates cardinal consistency, yet its CR is only 8%, and thusly the matrix is considered consistent. The authors generalize from this example and others that the 10% rule is easy to satisfy for small matrices (i.e. too lax and thereby admits inconsistent matrices) and is too hard to satisfy for larger ones (i.e. too restrictive and thereby excludes reasonably consistent matrices). [Aguaron and](#)

Moreno-Jimenez [8] as well as Bozoki and Rapcsak [13] state that Saaty [12] suggested changing the  $n = 3$  threshold to 5 %, and the  $n = 4$  threshold to 8%.

Holder [6] suggests that the linear scale used in REV is inappropriate and causes “inconsistency” in relative weight determination. He gives the following example to illustrate the point: “A is weakly more important than B (3 on Saaty's scale) and B is weakly more important than C (3 on Saaty's scale) imply that A is absolutely more important than C (9 on Saaty's scale)”. Holder also points out the negative consequences when assessing consistency due to the scale's discretization and limited upper bound value.

Murphy [7] gave a similar example to that of Golden and Wang's to show that the upper bound of the Saaty's Absolute 9-point scale forces choices outside the acceptable consistency range. The author states that the problem gets worse as the matrix size increases. Based on Saaty's 10% rule, for  $n = 3 - 9$ , Vargas [79] calculated upper limits for  $\lambda_{max}$  beyond which the PCMs would be inconsistent. Murphy concluded that in order to meet Vargas' criterion for consistency, for  $n = 3$  or 4, the average of adjacent elements in the PCM must be less than 5. For  $n$  larger than 4, Murphy recommends against using REV unless the average of adjacent elements in the PCM is less than 3.



[Genest and Rivest \[43\]](#) criticize the usage of the mean values of the RI distributions because the plots are highly skewed, and therefore the mean is not a representative parameter. The authors see the inclusion of RI in CR calculation as a positive contribution towards accounting for the concept of lower inconsistency limits when few elements are compared and higher allowable inconsistency when more elements are compared. Saaty had pointed out that this feature was “a theoretical confirmation of Miller’s psychological observation” [\[10\] \[80\]](#). Nonetheless, [Genest and Rivest \[43\]](#) insist that a statistical hypothesis testing approach is the recommended approach where such a feature is incorporated while accurately setting the consistency limits to accommodate the various decision making conditions

[Karapetrovic and Rosenbloom \[81\]](#) show several examples of quite reasonable matrices that are consistent with the views of the DM, yet they fail the standard CR recommended for REV. To remedy such situations, the authors recommend a quality control approach. The consistency indices of several PCMs are plotted on a range control chart to assess the DM’s consistency as a process. If the observations are out of control, the matrices are considered inconsistent. Otherwise, the DM’s choices are considered reasonable.

[Salo and Hämäläinen \[82\]](#) attribute the lack of accuracy using REV scale to the uneven weight of scale gradations. They point out, for example, that the difference in replacing

1 by 2 is 15 times the difference between 8 and 9. They also show that the upper bound of the REV scale causes the undesirable effect of restricting the range of the weight vector as follows:

$$w_{max} = \frac{M}{n + M + 1} \quad w_{min} = \frac{M}{M(n - 1) + 1} \quad (2.2)$$

where  $M$  is the upper scale bound and  $n$  is the number of comparison elements.

[Bana e Costa and Vansnick \[83\]](#) test the REV compliance with the Condition of Order Preservation (COP). This condition means the following: if, for example, 4 alternatives are judged such that  $A_1$  is preferred over  $A_2$  to a greater extent than  $A_3$  is preferred over  $A_4$ , then the weight vector should be such that:

1.  $(w_1 > w_2)$  and  $(w_3 > w_4) \cdots$  Preservation of order
2.  $(w_1 / w_2) > (w_3 / w_4) \cdots$  Preservation of intensity

The authors give three examples of PCMs, one of which is Saaty's own, where they all meet the CR requirement, yet they all violate COP. In the fourth example, the DM's choices make it impossible for the PCM to comply with COP. Yet again, CR for this matrix is well within the 10% CR rule. In 3 out of the 4 examples, the authors list the priority vectors obtained using a different prioritization procedure where COP is observed.

## 2.5 Other Consistency Measures

[Barzilai and Golany \[84\]](#) advocate the geometric mean method and claim it is the only acceptable method for multiplicative PCM's to simultaneously satisfy immunity to rank reversals, independence of problem description, independence of scale inversion, left-right eigenvector asymmetry, uniqueness, independence of order of operations and inter-level consistency. For additive PCM's, the authors suggest using the arithmetic mean method, for which they derive the following consistency measure:

$$\epsilon_{ij} = \frac{1}{n} \sum_{k=1}^n (a_{ij} + a_{jk} + a_{ki}) \quad (2.3)$$

where  $\epsilon_{ij}$  is the average inconsistency over all triplets with fixed  $i$  and  $j$ .

[Golden and Wang \[71\]](#) advocate using the Row Geometric Mean Method. They seek to develop a measure of consistency that is easy to use, is a function of matrix size, and has an intuitively appealing probability distribution. The formula for the Row Geometric Mean is:

$$g = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{bmatrix} = \begin{bmatrix} \sqrt[n]{C_{11}C_{12} \cdots C_{1N}} \\ \sqrt[n]{C_{21}C_{22} \cdots C_{2N}} \\ \vdots \\ \sqrt[n]{C_{N1}C_{N2} \cdots C_{NN}} \end{bmatrix} \quad (2.4)$$

where  $C_{ij}$  are the PCM elements.

The weight vector is normalized:

$$g_i^* = \frac{g_i}{\sum_i g_i} \quad (2.5)$$

Each column vector is normalized:

$$c_{ij}^* = \frac{c_{ij}}{\sum_i c_{ij}} \quad (2.6)$$

The consistency measure is:

$$G = \frac{1}{N} \sum_i \sum_j |c_{ij}^* - g_i^*| \quad (2.7)$$

In an analogous fashion, the authors suggest a consistency measure for REV as:

$$E = \frac{1}{n} \sum_i \sum_j |c_{ij}^* - e_i^*| \quad (2.8)$$

where  $e_i^*$  is the normalized eigenvector based on the principal eigenvalue.

The frequency histograms of  $G$ , generated from a simulation study using 1,000 matrices, were plotted. The distribution of  $G$  was approximately normal for  $n \geq 4$ . The Kolmogorov-Smirnov and the chi-squared tests were used to verify normality. It was observed that the normal distribution provided a good fit for  $n \geq 4$ . The authors obtained similar results for the distribution of  $E$ .

Ra [85] recommends HDM CS as a prioritization procedure but to measure consistency he recommends the Sum of Inverse Column Sums (SICS). The procedure is very simple, the columns of the PCM are summed, and then their reciprocals are added. SICS ranges in value from 0 to 1 with 1 indicating perfect consistency. Based on a simulation study using 1,000 matrices, Ra provides thresholds for SICS for  $n = 3 - 13$ . The “Standard” limits ensure very good consistency compliance, while the “Average” limits ensure ordinal consistency with minimal cardinal consistency violation. In a later effort (Kretchik and Ra [86]), SICS is presented as a consistency measure that is easy to use, is independent of the prioritization procedure, and is well bounded. Beta distribution is shown to be a good fit for SICS probability distribution.

**Table 2.3: Thresholds for the Sum of Inverse Column Sums (SICS) Method**

Thresholds for SICS		
n	Standard	Average
3	96.65%	85.74%
4	93.16%	75.89%
5	90.01%	67.36%
6	87.18%	59.91%
7	84.67%	54.92%
8	82.40%	51.48%
9	80.38%	48.38%
10	78.50%	46.38%
11	76.81%	44.80%
12	75.24%	42.70%
13	73.79%	41.20%

Jensen and Hicks [87] indicate that when considering tradeoffs regarding validity, reliability, consistency, and solution determinacy, being “finely cardinal” is not necessarily better or worse than being “coarsely ordinal”. Strictly for ordinal consistency, they propose to use Kendall’s Coefficient of Consistency, and provide computation formulas for the cases of with or without preference equivalence. The general formula is:

$$\zeta = \frac{D - d_0}{D} \quad (2.9)$$

where  $D$  is maximum number of circular triads for  $n$  items being compared, and  $d_0$  is the observed number of circular triads.

Koczkodaj [76] seeks to develop a consistency measure that is easy to interpret, allows easy selection of thresholds, and can link inconsistency to a particular element rather than an abstract value such as  $\lambda_{max}$ . His new consistency measure is computed among each triplet of the PCM elements. The formula for the consistency measure for a single triplet (a, b, c) is:

$$CM(a, b, c) = \min\left(\frac{1}{a}\left|a - \frac{b}{c}\right|, \frac{1}{b}\left|b - ac\right|, \frac{1}{c}\left|c - \frac{b}{a}\right|\right) \quad (2.10)$$

Bozoki and Rapcsak [13] extend the previous definition to the entire PCM (A for example):

$$CM(A) = \max \left\{ \min \left\{ \left| 1 - \frac{b}{ac} \right|, \left| 1 - \frac{ac}{b} \right| \right\} \text{ for each triplet } (a, b, c) \text{ in } A \right\} \quad (2.11)$$

The number of triplets in any  $n \times n$  matrix is:

$$\binom{n}{3} = \frac{n(n-1)(n-2)}{3!} \quad (2.12)$$

[Takeda \[88\]](#) developed the Measure of Consistency (MC) for the Row Geometric Mean

Method:

$$MC = \left[ \frac{1}{N(N-1)} \right] \sum_{\substack{i,j=1 \\ i \neq j}}^N a_{ij} c_{ij} \quad \text{and} \quad a_{ij} c_{ij} = \left[ \prod_{k=1}^N (a_{ij} a_{jk} a_{ki}) \right]^{\frac{1}{N}} \quad (2.13)$$

$MC \geq 1$ , for all reciprocal matrices

[Wedley \[89\]](#) suggested that for the sake of efficiency, instead of having to fill  $n(n-1)/2$

paired comparisons, a DM needs only to do  $n-1$  comparisons. The rest of the

comparisons are redundant and can be filled by a computer algorithm. Such a

procedure, in addition to achieving efficiency, will also build/maintain good ordinal

consistency. Wedley's suggested consistency measure for filling incomplete PCMs

(AVABDVCI: Average Absolute Deviation in Consistency Indexes) is:

$$AVABDVCI = \sum_{s=1}^{NR_s} \frac{ICI_s - ICI_{s-1}}{NR_s} \quad (2.14)$$

where  $ICI_s$  is the incomplete consistency index at step  $s$ , and  $NR_s$  is the number of redundant comparisons to this step.

[Takeda and Yu \[90\]](#) in the course of exploring the usefulness of using subsets of comparisons in PCMs to calculate the priority vector, developed the following consistency measure:

$$k = \left[ \prod_{(i,j) \in \mathcal{I}} a_{ij} \right]^{\frac{1}{n}} \quad (2.15)$$

where  $\mathcal{I}$  is the collection of the pair indices with which  $a_{ij}$  is described.

[Monsuur \[91\]](#) seeks to develop a consistency measure that is intrinsic, easily interpretable, scale independent, and can be adjusted to the decision situation.

Monsuur's measure is:

$$k = \frac{1}{2} \left( 1 - \lambda_{max} + \sqrt{[\lambda_{max} - 1]^2 + 4n} \right) \quad (2.16)$$

The author recommends an upper limit for the consistency measure of  $k \geq 0.9$ . For  $n = 3 - 9$  these thresholds closely match  $\lambda_{max}$  upper limits for consistent PCMs calculated by [Vargas \[79\]](#).

[Salo and Hämäläinen \[82\]](#) developed a scale-invariant Consistency Measure (CM) which is obtained through the equation:



$$CM = \sqrt{\frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n \left( a_{ij} - \log \frac{w_i}{w_j} \right)}{n(n-1)/2}} \quad (2.17)$$

[Barzilai \[92\]](#) develops consistency measure RC (Relative Consistency) for the multiplicative case based on his earlier measure for the additive case [\[84\]](#). First, the multiplicative PCM (M) is transformed to its “equivalent” additive matrix (A). Next the consistent components of A are computed:

$$C_A = (c_{ij}) = (w_i - w_j) \quad (2.18)$$

Finally, RC is computed:

$$RC(M) = \frac{\sum_{ij} c_{ij}^2}{\sum_{ij} a_{ij}^2} \quad (2.19)$$

[Shiraishi, et al. \[93\]](#) show that the characteristic polynomial of a PCM is:

$$P_A(\lambda) = \lambda^n + c_1 \lambda^{n-1} + \dots + c_{n-1} \lambda + c_n \quad (2.20)$$

They further define a new consistency measure “ $c_3$ ” as the coefficient of the characteristic polynomial:

$$c_3 = \sum \left\{ 2 - \left( \frac{a_{ij}a_{jk}}{a_{ik}} + \frac{a_{ik}}{a_{ij}a_{jk}} \right) \right\} \quad (2.21)$$

Crawford and Williams [78] developed a consistency measure for the Row Geometric Mean Method (RGMM). Aguaron and Moreno-Jimenez [8] formalized the measure calling it the Geometric Consistency Index (GCI), and provided the thresholds associated with it. The formula for GCI is:

$$GCI = \frac{2}{(n-1)(n-2)} \sum_{i < j} \left( \log(a_{ij}) - \log\left(\frac{w_i}{w_j}\right) \right)^2 \quad (2.22)$$

where  $e_{ij} = a_{ij}w_i/w_j$  is the error obtained when the ratio  $w_i/w_j$  is approximated by  $a_{ij}$ .

Because of GCI's independence of order, the authors established a relationship to CR in order to compute the thresholds:  $GCI = k(n)CR$ . The authors did a simulation study for  $n = 3 - 16$  using a sample of 100,000 matrices for each. Three sets of thresholds are provided at 4  $\alpha$  levels: for  $n = 3$ ,  $n = 4$ , and for  $n > 4$ .

Peláez and Lamata [94] seek to develop a Consistency Index (CI\*) that is easy to use, is a function of matrix size, and is applicable to other types of reciprocal matrices. For  $n = 3$ , the consistency measure is the determinant of the matrix. For  $n \geq 4$ , the consistency measure is the sum of determinants of all triplets (transitivities/triads) divided by the count of triplets. In other words, for  $n \geq 4$ , the consistency measure is the average determinant of triplets. To set limits for use of their consistency measure,

the authors carried out a simulation study of 100,000 matrices each for  $n = 3 - 9$ . The authors list the percentiles as thresholds for  $CI^*$

Given PCM  $M_{n \times n}$

$$CI_{N < 3}^* = 0, \quad CI_{N=3}^* = \det(M) \quad \text{and} \quad CI_{N > 3}^* = \frac{\det(\Gamma_M^*)}{NT_M} \quad (2.23)$$

where  $\Gamma_M^*$  is the collection of all triplets/transitivities, and  $NT_M$  is the number of triplets/transitivities.

[Gass and Rapcsák \[95\]](#) develop the Singular Value Decomposition (SVD) method as a prioritization procedure. They suggest the Frobenius norm of the difference between the original PCM and one formed by the SVD as an Inconsistency Measure (IM). The authors note that linking this measure to practical application and the DM's confidence still need to be developed. Their formula is:

$$IM = \sqrt{\sum_{i=1}^n \sum_{j=1}^n \left( a_{ij} - \frac{w_i}{w_j} \right)^2} \quad (2.24)$$

[Alonso and Lamata \[9\]](#) developed a statistical consistency acceptance criterion that is less restrictive than REV's, is linked to matrix size, is simpler, and has thresholds which are based on  $\alpha$  levels. Their consistency acceptance criterion is:

$$\lambda_{max} \leq n + \alpha(1.7699n - 4.3513) \quad (2.25)$$

Fedrizzi and Giove [96] develop a method for calculating the missing elements of an additive incomplete PCM. They do so by minimizing their measure of global inconsistency ( $\rho$ ). The measure is the mean value of the local consistency indices for all the possible triplets  $(x_i, x_j, x_k)$ :

$$\rho = \frac{\sum_{i < j < k} (r_{ij} - r_{ik} - r_{kj} + 0.5)^2}{\binom{n}{3}} \quad (2.26)$$

Stein and Mizzi [97] suggest using the Harmonic Consistency Index. Using a 500 sample of 4x4 randomly populated matrices, the authors try to show the correlation between HCI and REV's CI. The harmonic sum  $HM(s)$  is simply the sum of the inverse sum of the columns which is identical to what Ra [85], and Kretchik and Ra [86] have proposed years earlier. The authors state that HCI is well bounded and reaches its maximum value at the maximally intransitive matrix in a similar fashion to CI. The HCI's formula is:

$$HCI = \frac{[HM(s) - n](n + 1)}{n(n - 1)} \quad (2.27)$$

Fedrizzi and Brunelli [98] show that the further the pairwise judgments are from the neutral "indifference" position, the harder it is for the DM to achieve consistency, and vice versa. This phenomenon which the authors call "strength of preference effect" results in the DM with strong preference choices being penalized. The authors state that almost all consistency measures suffer from this shortcoming. To remedy this situation, they offer a new approach for assessing consistency which they call

“consistency equivalence classes”. By taking strength of preference effect into account, a class for either multiplicative or additive matrices is defined. The consistency level for a representative matrix from the class is calculated and is used as a “consistency level” for all the class members.

Čaklović [99] utilizes the Potential Method (PM), which represents pairwise comparisons through preference graphs to determine their results, to develop a new consistency measure. In PM, inconsistency is defined as the angle between the original preference flow and its consistent approximation. To set the thresholds for this consistency measure, the author performs 2 sets of simulations for  $n = 3 - 15$ . The first set uses normally-distributed randomly-generated perturbations, and the second uses uniformly-distributed randomly-generated perturbations. For  $n \geq 4$ , the consistency measure was found to closely follow a Gumbel distribution. The author listed the Gumbel distribution parameters for both sets of simulations along with the corresponding 5<sup>th</sup> percentiles as recommended upper thresholds.

Matteo, *et al.* [100] compared 2 pairs of the above mentioned consistency indices:  $CI^*$  from Peláez and Lamata [94] to  $c_3$  from Shiraishi, *et al.* [93], as well as  $GCI$  from Aguaron and Moreno-Jimenez [8] to  $\rho$  from Fedrizzi and Giove [96]. The authors prove proportionality between the first pair as well as the second. The authors concluded that

their results and similar ones should be used by researchers before embarking on developing new consistency measures in order to avoid duplication of effort.

Siraj, *et al.* [101] develop a new prioritization procedure based on all possible element combinations from tree spanning of pairwise comparisons. In the new method, Enumerating All Spanning Trees (EAST), the weight vector is composed of the average of individual weights computed for each tree. The consistency measure is the variance among the weight vector:

$$w = \frac{1}{\eta} \sum_{s=1}^{\eta} \tilde{w}(\tau_s) \quad \text{and} \quad \sigma^2 = \frac{1}{n(\eta - 1)} \sum_i \left( \sum_s (w_i - \tilde{w}_i(\tau_s))^2 \right) \quad (2.28)$$

where  $\eta = n^{n-2}$  is the number of spanning trees.

## 2.6 Analysis of Research Gaps

To establish the significance of the research topic and to identify the research gaps, several books and more than 180 journal articles have been reviewed. The table below summarizes the research gaps identified.

**Table 2.4: List of Research Gaps**

No	Article	Critical Issue Discussed	Research Gaps
1	<a href="#">Barzilai and Golany [84]</a>	Provide consistency measure for additive Pairwise Comparison Matrices (PCMs)	The multiplicative case is not addressed, no thresholds for the measure are provided, and the measure is not linked to number of elements or $\alpha$ levels. The results are not in ratio scale.
2	<a href="#">Golden and Wang [71]</a>	Provide consistency measure for Row Geometric Mean Method (RGMM)	The measure is applicable only to RGMM. Although the measure is linked to number of elements, thresholds and $\alpha$ levels would be hard to establish due to the rough fit to normal distribution.
3	<a href="#">Ra [85]</a> , <a href="#">Kretchik and Ra [86]</a>	Provide consistency measure for HDM's Constant Sum (CS)	The fit of the measure to a beta distribution is quite rough. This made linking it to number of elements weak and consequently multiple thresholds and $\alpha$ levels were not established. No theoretical justification for the measure is given, and therefore its interpretation is unknown. Finally, the upper bound for the measure is unproven particularly for higher order matrices.
4	<a href="#">Jensen and Hicks [87]</a>	Provide ordinal consistency measure for pairwise comparisons	Measures only ordinal consistency. No thresholds for the measure are provided, and the measure is not linked to number of elements or $\alpha$ levels.

No	Article	Critical Issue Discussed	Research Gaps
5	<a href="#">Koczkodaj [76]</a>	Provide consistency measure for pairwise comparison matrices (PCMs)	Good measure with the added benefit of locating inconsistency within a triplet. Few recommended thresholds have been established for low order matrices. Extending these to higher orders is yet to be done. Thresholds are not linked to $\alpha$ levels.
6	<a href="#">Takeda [88]</a>	Provide consistency measure for RGMM	The measure is applicable only to RGMM. No thresholds for the measure are provided, and the measure is not linked to number of elements or $\alpha$ levels.
7	<a href="#">Wedley [89]</a>	Provide a per step consistency check for use while filling incomplete matrices	The measure is specifically for filling incomplete matrices. No thresholds for the measure are provided, and the measure is not linked to number of elements or $\alpha$ levels.
8	<a href="#">Takeda and Yu [90]</a>	Provide a consistency measure for a subset of a Pairwise Comparison Matrix (PCM)	The measure is specifically for a subset of a PCM. No thresholds for the measure are provided, and the measure is not linked to number of elements or $\alpha$ levels.
9	<a href="#">Monsuur [91]</a>	Provide an intrinsic consistency measure that is scale independent	The measure is linked to the abstract quantity of maximum eigenvalue. No statistically based thresholds or corresponding $\alpha$ levels are provided.
10	<a href="#">Salo and Hämäläinen [82]</a>	Provide consistency measure that is scale invariant	The measure is more suitable for distance-minimizing methods. The measure is not linked to the matrix order and no statistically based thresholds or corresponding $\alpha$ levels are provided.



No	Article	Critical Issue Discussed	Research Gaps
11	<a href="#">Barzilai [92]</a>	Provide consistency measure that ensures immunity to rank reversal, independence of problem description, independence of scale inversion, left-right eigenvector asymmetry, uniqueness, independence of order of operations and inter-level consistency	The measure has many advantages. However, it is cumbersome to calculate (involves transforming multiplicative PCMs to their additive equivalents), is unbounded, and lacks statistically based thresholds and their corresponding $\alpha$ levels.
12	<a href="#">Shiraishi, et al. [93]</a>	Provide consistency measure for positive reciprocal matrices	The measure is not linked to the matrix order and no statistically based thresholds or corresponding $\alpha$ levels are provided.
13	<a href="#">Crawford and Williams [78]</a> , <a href="#">Aguaron and Moreno-Jimenez [8]</a>	Provide consistency measure for RGMM and provide thresholds for the measure	The measure is applicable only to RGMM. Because of GCI's independence of order, the thresholds were approximated by establishing a relationship to CR. The thresholds are provided for $n = 3, 4$ , and all matrices $> 4$ . This makes the measure's link to the number of elements quite weak. Only 4 $\alpha$ levels were given.
14	<a href="#">Peláez and Lamata [94]</a>	Provide a consistency index that is easy to use, is a function of matrix size, and is applicable to other types of reciprocal matrices	The measure is a function of matrix size, has statistically based thresholds, and corresponding $\alpha$ levels. However, the thresholds and the $\alpha$ levels are for Saaty's scale only, the measure is mathematical and its applicability to stochastic methods such as judgment quantification in HDM is questionable. Is this measure proportional to HDM's inconsistency measure? Will it work in concert with HDM's statistical prioritization procedure?
15	<a href="#">Gass and Rapcsák [95]</a>	Provide consistency measure for Singular Value Decomposition (SVD) method	The measure is applicable only to SVD. No thresholds for the measure are provided, and the measure is not linked to number of elements or $\alpha$ levels.
16	<a href="#">Alonso and Lamata [9]</a>	Provide consistency measure for REV method	The measure is a function of matrix size, has statistically based thresholds, and corresponding $\alpha$ levels. However, it is applicable only to REV.

No	Article	Critical Issue Discussed	Research Gaps
17	<a href="#">Fedrizzi and Giove [96]</a>	Provide consistency measure for additive PCMs	The measure is applicable only to additive PCMs. No thresholds for the measure are provided, and the measure is not linked to number of elements or $\alpha$ levels.
18	<a href="#">Stein and Mizzi [97]</a>	Provide consistency measure for PCMs	The measure is theoretically applicable to all PCMs. No thresholds for the measure are provided, and the measure is not linked to number of elements or $\alpha$ levels.
19	<a href="#">Fedrizzi and Brunelli [98]</a>	Provide a consistency approach that takes into account “strength of preference effect”	A consistency approach rather than a measure.
20	<a href="#">Čaklović [99]</a>	Provide consistency measure for the Potential Method	The measure is a function of matrix size, has statistically based thresholds, and corresponding $\alpha$ levels. However, it is only applicable to the Potential Method.
21	<a href="#">Siraj, et al. [101]</a>	Provide consistency measure for the method of Enumerating All Spanning Trees (EAST)	The method is very similar to HDM’s: it is stochastic rather than deterministic, and the weight vector is the average of variable weights which are computed for many “orientations”. The consistency measure however is applicable only to “EAST”, is not linked to the matrix order and no statistically based thresholds or corresponding $\alpha$ levels are provided.

There is consensus in the literature on the importance of defining, understanding, controlling, and improving consistency in AHP in order to build reliability, confidence and meaningfulness in the entire process of AHP decision making [\[3, 12, 71, 76, 82, 91, 92, 94, 97\]](#). Considerable research effort on achieving these goals for AHP consistency spans the 3 decades since the introduction of this decision-making tool. This proves the significance of this topic for both researchers and practitioners.

As a variant of AHP, HDM's procedure for judgment quantification, which was developed by D.F. Kocaoglu, defines inconsistency as follows:

$$Inconsistency = \frac{1}{n} \sum_{i=1}^n \sigma_i \quad (2.29)$$

where  $\sigma_i$  is the standard deviation of the mean of  $n$  factorial normalized relative values for the  $i^{th}$  decision element, and  $n$  is the number of decision elements.

The acceptable limit is 0.1. It does not vary with number of elements, and is not linked to a  $\alpha$  levels.

In summary, many of the prioritization procedures lack an inconsistency measure, and many of the ones that do provide inconsistency measures have global limits defined without considering the number of elements involved or the  $\alpha$  levels required [8, 9].

Clearly, this major research gap presents an opportunity to complete the development of this important metric.

Wide-spread research [8, 9, 71, 86, 99, 102] indicate that a statistical approach built on the estimated probability distribution of the inconsistency parameter is the way to achieve the desirable inconsistency properties of

- a) Being a function of the number of elements
- b) Having limits linked to  $\alpha$  levels

The research presented in this dissertation does this for the Hierarchical Decision Model (HDM) procedure.

## 2.7 Statistical Approach to Establishing Consistency Thresholds

Evidence from current research shows there is consensus that a simulation-based statistical approach to establishing consistency thresholds is the best justified and most adopted by researchers in this field. Following is a chronological review of research on this issue.

[Vargas \[79\]](#) conducted the earliest research on the validity of thresholds related to REV's consistency measure. His study was based on a simulation that used 500 matrices each for  $n = 3 - 9$ . The author listed the mean of the consistency measure for each matrix order, the consistency magnitude at  $\alpha = 5\%$ , and the corresponding  $\lambda_{max}$  at this confidence level. The author also carried out curve-fitting analysis and concluded that REV's consistency measure followed a Dirichlet distribution. To measure the goodness of fit, the author used the Kolmogorov-Smirnov test (KS-test). The matrices used in the simulation were randomly filled using discrete uniform distribution.

[Lane and Verdini \[72\]](#) carried out a simulation study to establish the distribution of REV's consistency measure and study its implications to the thresholds set by Saaty (the 10% rule). The study used the full enumeration of matrices for  $n = 3$  (4,913), and 2,500 matrices for  $n = 4 - 10, 12, 14, 16, 20$ , and 24. The authors state that their statistical hypothesis-testing approach will enable simulation and measurement of DM's

randomness. In turn, through establishing the distribution of the consistency measure, users can choose a consistency limit which corresponds to an  $\alpha$  level suitable for their needs. The matrices used in the simulation were randomly filled.

[Golden and Wang \[71\]](#) carried out a simulation study to investigate the validity of the recommended thresholds related to REV's consistency measure. The authors used samples of 1,000 matrices each for  $n = 3 - 15$ . Based on their statistical simulation they recommended a new consistency measure for which they also did statistical simulation to determine appropriate thresholds. They recommended using a consistency threshold at  $\alpha = 33\%$ . The authors also carried out curve-fitting analysis and concluded that their new consistency measure roughly followed a normal distribution. To measure the goodness of fit, they used the Kolmogorov-Smirnov test (KS-test). The matrices used in the simulation were not randomly but according to an algorithm that assumes and simulates a reasonable level of consistency.

[Dodd, et al. \[102\]](#) carried out a simulation study to investigate the validity of the recommended thresholds related to REV's consistency measure. The authors used samples of 1,000 matrices each for  $n = 4 - 10$ . Based on their statistical simulation they recommended the adoption of new consistency thresholds which are based on the probability distribution of the consistency measure and correspond to  $\alpha$  levels. The

permitted CI thresholds corresponding to  $\alpha$  levels of 0.1%, 0.5%, and 1%-5% were reported. The matrices used in the simulation were randomly filled.

[Noble and Sanchez \[41\]](#) carried out a simulation study to establish the distribution of entropy as a measure of inconsistency. Samples of 1,000 matrices each for  $n = 3 - 15$  were used. The mean random entropy value for each order was reported. The Kolmogorov-Smirnov Lilliefors test for normality was used to show that the entropy distribution was normal. The authors started with randomly filling the first row of a matrix. "The rest of the entries  $X_{ij}$  of the matrix were obtained by direct computation, to force the matrix to be totally consistent" [\[41\]](#). These randomly-initialized consistent matrices were then randomly perturbed.

[Kretchik and Ra \[86\]](#) carried out a simulation study to determine appropriate thresholds for Ra's [\[85\]](#) Sum of the Inverse Column Sums (SICS) proposed inconsistency measure. Many simulation runs were done with a maximum sample size of 60,000. Curve-fitting analysis was conducted and the authors concluded that the new consistency measure followed a beta distribution. The matrices used in the simulation were randomly filled.

[Aguaron and Moreno-Jimenez \[8\]](#) carried out a simulation study to determine appropriate thresholds for the Geometric Consistency Index (GCI) used in conjunction with the Row Geometric Mean Method both of which were first proposed by

Crawford and Williams [78]. The authors used 100,000 matrices each for  $n = 3 - 15$ . The permitted thresholds for  $\alpha$  levels of 1%, 5%, 10%, and 15% were reported. The matrices used in the simulation were randomly filled.

Peláez and Lamata [94] carried out a simulation study to determine appropriate thresholds for their new proposed consistency index  $CI^*$ . The authors used 100,000 matrices each for  $n = 3 - 9$ . The permitted thresholds for  $CI^*$  at  $\alpha$  levels of 1% and 5% – 50% in increments of 5% were reported. The matrices used in the simulation were randomly filled.

Alonso and Lamata [9] aimed to develop a new consistency measure for REV that is a function of PCM order and is tied to an  $\alpha$  level. They carried out a simulation study to determine appropriate thresholds for the proposed measure. Two sets of samples were used, the first was made up of 100,000 matrices and the second of 500,000 matrices each for  $n = 3 - 15$ . The authors reported they found no differences in results between the two sets. Being a function of the PCM size and level of  $\alpha$  was built into the formula the authors derived for the new consistency criterion:

$$\lambda_{max} \leq n + \alpha(1.7699n + 4.3513) \quad (2.30)$$

Therefore, there was no need to separately list thresholds. The matrices used in the simulation were randomly filled.



Bozoki and Rapcsak [13] carried out a simulation study to compare Saaty's CR to Koczkodaj's Consistency Measure (CM). Samples of 10,000,000 matrices each were used for  $n = 3 - 10$ . The results for CR were plotted and the  $\alpha$  levels identified. For  $n = 6 - 10$ , the curves show Saaty's recommended 10% rule well outside to the left of the distributions. The matrices used in the simulation were randomly filled.

Čaklović [99] developed a consistency measure for the Potential Method and carried out a simulation study to determine its admissible thresholds. Samples of 100,000 matrices each for  $n = 3 - 9$  were used. The thresholds for the measure were listed at 5%  $\alpha$  level. The author also carried out curve-fitting analysis and concluded that the consistency measure closely followed a Gumbel distribution. The formula for Gumbel distribution was given along with its parameters for the consistency measure for  $n = 3 - 9$ . The matrices used in the simulation were randomly filled.

## 2.8 Summary of Literature Review

The credibility of an AHP/HDM model hinges on having sufficient compliance with consistency rules. Nonetheless, bias in judgment, human error, inaccuracy, and inexperience can all contribute to the violation of consistency rules. If all these factors cause gross inconsistency, the AHP/HDM model will be rendered useless. Because of this reason, a great deal of research effort has been devoted to studying the subject of consistency in pairwise comparison [Kou, \*et al.\* \[103\]](#).

It is important to define thresholds that are based on the size of the decision problem and are tied to statistical  $\alpha$  levels. To determine such thresholds, research in the field clearly shows that simulation of randomly generated input matrices to obtain the distribution of the consistency measure is a widely used method. An added benefit to this approach is to perform curve-fitting analysis for identifying the known distribution that the inconsistency measure closely follows. That way, users can directly use the distribution's formula along with specific size parameters to determine the consistency threshold corresponding to their choice of  $\alpha$  level.

## 3 RESEARCH

### 3.1 Research Approach

This PhD research has developed a method for analyzing decision inconsistencies using the HDM's judgment quantification method in response to the key gap that has been identified in the literature. The research question is: How can HDM's consistency thresholds be defined to comply with the requirements of:

1. Being a function of the size of the decision problem.
2. Being subjected to hypothesis testing.
3. Being defined as a distribution.
4. Being linked to  $\alpha$  levels.

The literature review shows that the method of choice among researchers for defining consistency thresholds with the above desired properties is through computer simulation of randomly generated inputs into the judgment quantification methods.

HDM inconsistency is defined in this research as the square root of the sum of variances (RSV) of the means of  $n$  variables calculated in  $n$  factorial orientations:

$$RSV = \sqrt{\sum_{i=1}^n \sigma_i^2} \quad (3.1)$$

It is a modified version of the current inconsistency measure used in HDM, which is

$$Inconsistency = \frac{1}{n} \sum_{i=1}^n \sigma_i \quad (3.2)$$

This modification was necessary because the numerical values for the current measure were very small and the precision was being lost when large numbers of randomly generated input matrices were analyzed.

Below is the simulation procedure used for defining the consistency thresholds for HDM's judgment quantification method:

1. Setup input data structure: This involves building Matrix "A" which is an  $n \times n$  matrix.
2. Fill in the data structure: This will be done by populating either side of the left diagonal of Matrix "A" with randomly generated numbers in the range of 1 – 99. The other half of the matrix will be filled with the 100-compliment of the mirror positions on the other diagonal side.
3. Perform necessary calculations: This will involve building matrices "B" and "C", defining the  $n$  factorial orientations for all the elements, computing elements values for all orientations using the direct and indirect ratios derived from Matrix "C", calculating the standard deviation per element from all orientations, and finally computing the average standard deviation as the mean of inconsistency.

4. Store results: Save the results from each run.
5. Repeat the above steps: The above represents the computations for one set of simulated judgment inputs. The process will be repeated for 100,000 sets of input data because initial testing shows that stability is reached well before that level. Also, literature shows that there is no statistically significant difference in repeating the simulation beyond 100,000 cycles under any condition.
6. Analyze the results: This involves plotting the sample's histogram, determining the sample's statistical parameters such as the minimum, maximum, mean, percentiles, cumulative distribution function (CDF), and Quantile function.
7. Perform curve fitting and test goodness-of-fit (GOF) using the Kolmogorov–Smirnov test (K–S test).

## 3.2 Research Procedure

### 3.2.1 Sample Sizes

Sample sizes as small as 500 and as large as 10,000,000 have been used in the simulation studies for inconsistencies in other judgment quantification methods. Larger sample sizes do not necessarily mean better or more reliable simulation results. For example, [Aguaron and Moreno-Jimenez \[8\]](#) and [Alonso and Lamata \[9\]](#) used 100,000 while [Bozoki and Rapcsak \[13\]](#) used 10,000,000 but all have obtained almost identical results. [Golden and Wang \[71\]](#) used a relatively small sample of 1,000, yet their results differ by an average of 0.38%, 0.43%, 0.57% from these studies respectively. Finally, [Alonso and Lamata \[9\]](#) used two sets of sample sizes, one with 100,000 and the other with 500,000. They found no difference between the results of the two sets.

As indicated in the requirements of the solution, the results should be fitted to standard probability distributions. This is an important step towards making the results much more valuable in a practical sense. Requiring users of HDM to maintain a multitude of detailed distribution is cumbersome and somewhat impractical. Fitting the results to a known probability distribution would allow for a great deal of flexibility and portability. All the users have to know are merely the parameters for the fitted distribution related the number of elements they are using. Using this information, the users can easily use either:

- The standard equation for the cumulative distribution function (CDF) to calculate the  $\alpha$  level for a given magnitude of inconsistency, or
- The standard equation for the Quantile function (inverse of the cumulative distribution function) to calculate the maximum allowed magnitude of inconsistency given a specific/desired  $\alpha$  level.

In statistical literature, the Kolmogorov-Smirnov test (K-S test) is a widely used for goodness of fit (GOF). The K-S test is non-parametric in the sense that it compares the CDF of the empirical distribution obtained through simulation against the CDF of a hypothesized probability distribution without basing its results on the statistical parameters such as the means, maxima or the minima of the two distributions. Alas, the advantage of the K-S test can have unintended consequences. The test is quite sensitive to small differences between the empirical and hypothesized probability distributions. Large samples intensify this effect. [Gibbons \[104\]](#) indicates that for large samples, the test will almost always reject the null hypothesis (reject the empirical distribution as a bad fit to the hypothesized probability distribution). [Law and Kelton \[105\]](#) state that such a behavior “is an unfortunate property of these tests since it is usually sufficient to have a distribution that is nearly correct”. Therefore, care must be taken in order to strike a balance between the two competing priorities: a large enough sample to ensure accuracy and a small enough sample that will give satisfactory goodness of fit results.

According to [Gibbons \[104\]](#), the calculation of N, the minimum sample size required for the K-S test, must satisfy:

$$P(D < d) = 1 - \alpha \quad (3.3)$$

where D is the Kolmogorov-Smirnov test statistic, d is maximum allowed error, and  $1 - \alpha$  is the selected probability. For 99% precision and 95% confidence the minimum sample size is 1063 [\[104\]](#). For 99% precision and 99% confidence the minimum sample size is 6642.

### 3.2.2 High Quality Random Number Generator

The random number generator (RNG) provided with the C++ package is unfortunately a linear congruential generator (LCG). Such generators should not be used particularly in simulation because they suffer, among other defects, from serial correlation between successively generated numbers as well as short periods which could cause unacceptable repetitions [\[106, 107\]](#).

Instead of the built in RNG, the solution uses a C++ version of the Mersenne Twister which is a high quality RNG [\[108\]](#) with many desirable statistical properties [\[109\]](#) including a large period ( $2^{19937} - 1$ ) [\[110\]](#) and negligible correlation [\[111\]](#). The Mersenne Twister is most suitable for our study because it was designed with statistical simulation in mind [\[112\]](#). All these advantages made the Mersenne Twister the default



choice for widespread research software packages such as R, Matlab, and Swarm Simulation Systems [\[113\]](#).

### **3.2.3 Reducing Computational Intensity**

HDM's judgment quantification method is a statistically-based prioritization procedure. The weight/ranking of various alternatives is calculated as the average of their individual values from their factorial permutations. The HDM's inconsistency measure is based on the standard deviation among the individual values. The computational intensity of the calculations increases dramatically as the number of alternatives increases. For example if 10 elements are considered, then at least 36,288,000 calculations per data point would have to be performed. This does not include the generation of random numbers or the complex matrix operations. To illustrate this dilemma, a 15 hour experiment was conducted to help estimate the time required for 100,000 samples for 10 elements. The extrapolated estimate was an astounding 624 days. The experiment was done on a fast Intel Core i7-4790K machine running at 4.0 GHz with 32 Gigabyte of memory.

The first part of the solution was to use SIMD-oriented Fast Mersenne Twister (SFMT). This is a new variant of the original Mersenne Twister designed with recent parallelism of modern CPUs, such as multi-stage pipelining and SIMD (single instruction multiple

data allowing 128-bit integer generation). This RNG is supposed to be twice faster than the original version.

### **3.2.4 Programming Language and User Interface**

The implementation research of this dissertation was carried out in three phases. The objective of the first one was to test the proposal's theoretical assumptions and produce basic results upon which to design and build the next one. The second one involved a great deal of research to modify and strengthen the implementation as suggested by the previous results. The research of the third and final one provided the key solution to make the implementation reliable, robust, and practical.

A significant conclusion of the first stage of development was the need for substantial improvement in computational efficiency. The simulation code was taking too long to produce sufficiently large samples and this issue needed to be resolved. The first step in this endeavor was to identify performance bottlenecks and other areas where improvements were needed.

The language used in the first research stage was Visual Basic .NET (VB.NET). This was not a good choice for efficiency. In addition, despite the convenience provided by the user-friendly Windows interface, the overhead cost in terms of time associated with the interface turned it into a major obstacle.

There is almost consensus that for performance, efficiency and speed of execution, C++ is the language of choice. Compared to other .NET alternatives, C++ performs much better in [114]:

- Numeric calculation even when compared to C#
- Memory management
- Operating system interface and access

Other languages to consider such as Java offer considerably slower performance than C++ even when using appropriate optimizations and regardless of the operating system hosting the comparison tests [115].

Using Open Source languages such as R was not a viable option due to its extreme slowness: 243 – 282 times slower than C++ for compiled R code, and 475 – 491 times slower than C++ for interpreted R code [116].

In conclusion, C++ is the confirmed choice for speed and efficiency over any other alternative [116].

The success of the simulation program depended on performance and speed of execution. To achieve this end, the above clearly proved that C++ was the tool of choice. Furthermore, using a console application that is driven by command line options in C++ seemed the appropriate choice under the circumstances.

### **3.2.5 Procedural Efficiencies in Judgment Quantification Method**

In order to evaluate each of the variables in an orientation, the current procedure for HDM judgment quantification method calls for the computation of a matrix (Matrix C) for every orientation. To process the entire orientation table would require the construction of  $n$  factorial ( $n!$ ) matrices each of which is  $n$  by  $n-1$  in size. However, upon closer examination, it was observed that these matrices were not unique and that there was a great deal of computational redundancy. In fact, the number of unique columns shared by all these matrices was  $n \times (n-1)$ . The procedure was modified to build a single matrix that is  $n$  by  $(n \times (n-1))$  from which all the relevant ratios would be derived. This modification made the procedure significantly more efficient.

### **3.2.6 The Need for a Leap in Efficiency**

Despite the implementation of the above improvement measures, the gains in efficiency and speed of execution were still not enough to make the simulation program practical. This was true for large numbers of variables ( $n \geq 8$ ) and particularly so for  $n \geq 10$ . It was clear that the computations associated with an exhaustive solution using full enumeration of orientations, for large number of variables, were consuming too much time.

The issue of excessive computational burden due to exhaustive enumeration can often make solutions infeasible [117, 118]. This drawback of exhaustive enumeration limits its use to small problems [119]. Fortunately, such a disadvantage could be overcome through sampling, which usually produces results very close to those of exhaustive enumeration, a fact that has been simply verified through various empirical studies in a range of fields [120, 121].

The first implementation of this approach was done through randomly sampling the full orientation table. The preliminary results were quite encouraging. Nonetheless, there were two problems with the solution which prevented its adoption:

The minimum accuracy was unpredictable, and therefore could not be guaranteed. Increasing the sample size did not necessarily resolve the issue of minimum accuracy.

Upon further research, it was observed that HDM's lexicographically-ordered rows of the fully enumerated orientation table formed subgroups each starting with a different element. The number of subgroups equals the number of elements being evaluated. A more important observation was related to the values of elements within a subgroup versus those across subgroups. The values of elements obtained within each subgroup were a lot closer in magnitude to each other in contrast to the values of the same elements obtained within other subgroups.

The organization of the lexicographically-ordered orientation table with its non-overlapping subgroups seemed to fit the classic definition of homogenous subgroups and their analogous strata [122]. In this case, the stratifying variable was the first element in an orientation [123]. Most importantly, it seemed that the key to a stable and accurate solution was the implementation of stratified rather than random sampling. This means the sampling strategy required the samples to be “stratified” or dispersed evenly among the homogenous subgroups which would address the problem of unbalanced representation from which the random sampling suffered. This is a sampling technique that is widely used and offers flexibility and power [124].

In random sampling, which was used first, the homogenous subgroups were not equally represented. The random samples often overrepresented some subgroups while underrepresenting others. This explains the problem of mixed accuracy in random sampling.

### **3.2.7 Producing a Single Permutation Given its Index**

The implementation of stratified sampling required an efficient mapping function designed to generate only those specific orientations that were part of the sample. Producing all permutations iteratively to select a small subset seems counterproductive.

The Lehmer code [125] offers a systematic method to uniquely represent a permutation. This is an example of an inversion vector [126], a unique sequence of numbers from which the unique corresponding permutation can be reconstructed [127]. According to Roberto and Fanja [128], Lehmer's code-producing function  $l$ , which is bijective (gives unique one-to-one correspondence), acts upon a permutation  $\pi$  of length  $n$  to produce the sequence:

$$l(\pi) = (l_1(\pi) \cdots l_i(\pi) \cdots l_n(\pi))$$

where  $l_i(\pi)$  is the number of elements in the set  $\{j > i : \pi(j) < \pi(i)\}$  [129]

Pesko [130] listed the above and added the expression for the reconstruction of the permutation as follows:

$$\pi(k) = N_k[l_k(\pi) + 1], \quad \text{where } N_k = \{\pi(1), \pi(2), \dots, \pi(k-1)\}$$

The first half of table 3.1 below shows the construction of the Lehmer code of an 8-element permutation. The second half shows how the permutation can be recovered using its Lehmer code.

Original Permutation	Ordered Permutation Set	Lehmer Code for Indicated Element
{ 6 1 4 3 8 5 7 2 }	{ 1 2 3 4 5 <b>6</b> 7 8 }	5
{ 1 4 3 8 5 7 2 }	{ <b>1</b> 2 3 4 5 7 8 }	0
{ 4 3 8 5 7 2 }	{ 2 3 <b>4</b> 5 7 8 }	2
{ 3 8 5 7 2 }	{ 2 <b>3</b> 5 7 8 }	1
{ 8 5 7 2 }	{ 2 5 <b>7</b> 8 }	3
{ 5 7 2 }	{ 2 <b>5</b> 7 }	1
{ 7 2 }	{ 2 <b>7</b> }	1
{ 2 }	{ <b>2</b> }	0
Reconstructed Permutation	Ordered Permutation Set	Lehmer Code for Indicated Element
	{ 1 2 3 4 5 6 7 8 }	{ 5 0 2 1 3 1 1 0 }
{ 6 }	{ 1 2 3 4 5 <b>6</b> 7 8 }	{ <b>5</b> 0 2 1 3 1 1 0 }
{ 6 1 }	{ <b>1</b> 2 3 4 5 <b>6</b> 7 8 }	{ <b>5</b> <b>0</b> 2 1 3 1 1 0 }
{ 6 1 4 }	{ <b>1</b> 2 3 4 5 <b>6</b> 7 8 }	{ <b>5</b> <b>0</b> <b>2</b> 1 3 1 1 0 }
{ 6 1 4 3 }	{ <b>1</b> 2 <b>3</b> 4 5 <b>6</b> 7 8 }	{ <b>5</b> <b>0</b> <b>2</b> <b>1</b> 3 1 1 0 }
{ 6 1 4 3 8 }	{ <b>1</b> 2 <b>3</b> 4 5 <b>6</b> 7 <b>8</b> }	{ <b>5</b> <b>0</b> <b>2</b> <b>1</b> <b>3</b> 1 1 0 }
{ 6 1 4 3 8 5 }	{ <b>1</b> 2 <b>3</b> 4 5 <b>6</b> 7 <b>8</b> }	{ <b>5</b> <b>0</b> <b>2</b> <b>1</b> <b>3</b> <b>1</b> 1 0 }
{ 6 1 4 3 8 5 7 }	{ <b>1</b> 2 <b>3</b> 4 5 <b>6</b> 7 <b>8</b> }	{ <b>5</b> <b>0</b> <b>2</b> <b>1</b> <b>3</b> <b>1</b> <b>1</b> 0 }
{ 6 1 4 3 8 5 7 2 }	{ <b>1</b> <b>2</b> <b>3</b> 4 5 <b>6</b> 7 <b>8</b> }	{ <b>5</b> <b>0</b> <b>2</b> <b>1</b> <b>3</b> <b>1</b> <b>1</b> <b>0</b> }

**Table 3.1: Construction of Lehmer Code of a Permutation and the Permutation's Recovery from its Lehmer Code**



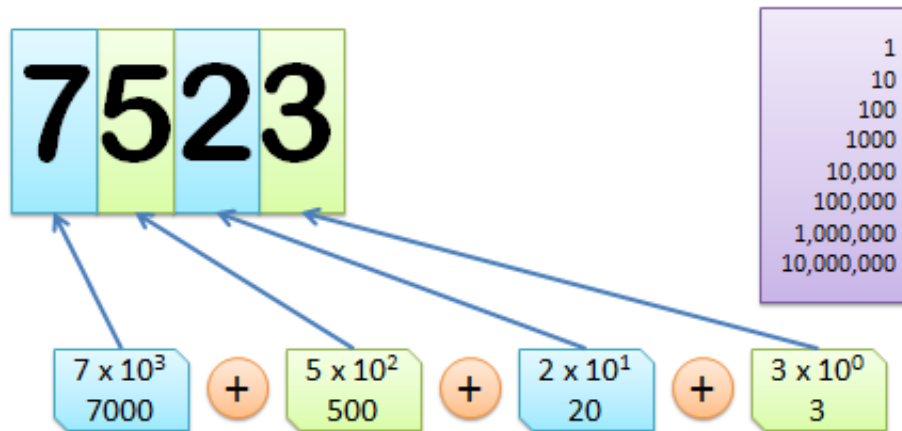
As important as Lehmer's discovery was, it left more to be desired. Lehmer showed that given a permutation, his code could be used to uniquely represent it and then reconstruct it. Nonetheless, the question became: what to do in the absence of the permutation itself? How can the code be generated given the permutation's lexicographic index?

[Knuth \[131\]](#), who cited [Hall \[132\]](#) rather than Lehmer when discussing unique permutation codes, provided the crucial answer to the above question by which the circle could be completed. Simply put, the factoradic representation of a permutation's lexicographic index is the permutation's Lehmer code. This revelation is powerful indeed.

The factoradic, or the factorial numbering system, uses a varying base (mixed radix) which is comprised of the factorial values any number  $n \in (1, 2, \dots, n)$ . For a dimension  $n$ , the factoradic can be used to uniquely represent each number in the range (0 to  $n! - 1$ ) [\[133\]](#). [Samarghandi, et al. \[134\]](#) cite [Knuth \[131\]](#) attributing the uniqueness of representation of the factoradic to the identity which states "the sum of consecutive factorials multiplied by their index is always the next factorial minus one" [\[134\]](#)

$$\sum_{i=0}^n i \times i! = (n + 1)! - 1 \quad (3.4)$$

## Decimal (Base 10)



## Factoradic (Base Factorial)

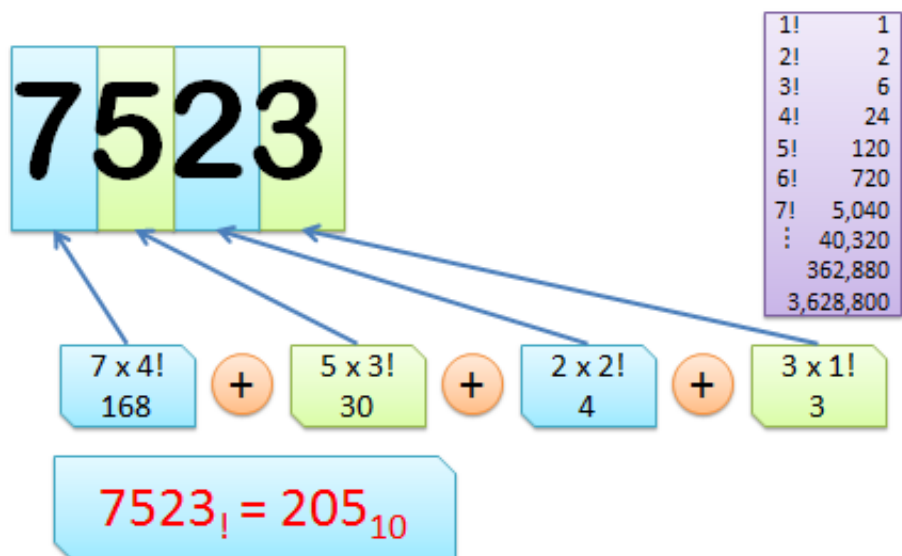


Figure 3: Decimal vs. Factoradic Representation of Numbers

### 3.2.8 Implementing the Stratified Sampling

The pair of these mathematical methods enabled the stratified sampling to be implemented with outstanding speed and efficiency. The solution steps are:

1. Determine an appropriate sample size
2. Slightly adjust the sample size so it is evenly divisible by  $n$  which is the number of homogeneous subgroups (strata).
3. The division of adjusted sample size by  $n$  yields the size of stratified sampling interval which should enable evenly dispersed sampling among the strata.
4. The beginning address of each sampling interval is converted from its decimal form to its factoradic equivalent.
5. The permutation is reconstructed from its Lehmer code (factoradic) and is used to fill the abridged orientation table.
6. The reduced orientation table is used for calculations

### 3.2.9 Testing Stratified Sampling

In order to assess the effectiveness and accuracy of the stratified sampling solution, an experiment was conducted on 100 ten by ten input matrices. The results of the full enumeration exhaustive method were compared to those of stratified sampling. With a tiny sample of 10,000 which is merely 0.28% of the full sample space of 3,628,800, the

stratified sampling method produced results that matched those from the exhaustive method to at least the fourth decimal digit. Repeated experiments confirmed the same minimum accuracy for the same stratified sample size. This verification confirms that not only is the stratified sampling method accurate, fast, and efficient, but also very stable.

## 4 RESEARCH RESULTS

### 4.1 Design of Simulation Study

The original objective of the research included the determination of inconsistency thresholds for decision variables in the range of  $n = 3 - 10$ . This was expanded to cover the range of  $n = 3 - 12$ .

The simulation study carried out for this research was based on a set of three samples: 10k, 100k, and 500k. The sample size choices were made as follows:

- 10k sample: This is a size that well exceeds the minimum sample size of 6642 for 99% precision and 99% confidence required for the K-S GOF test as indicated by [Gibbons \[104\]](#) which is discussed in section 3.2.1. While meeting and exceeding this requirement, the 10k sample size is not too large as to cause the K-S GOF test to produce poor fit results as described by [Law and Kelton \[105\]](#). After verification with the empirical data from the three sample sets, the distribution fit results and related CDF and Quantile functions from this sample will be used as the recommended tools for consistency threshold calculations for HDM.
- 100k sample: This is a large sample size as to allow for greater accuracy and good precision when comparing empirical and fit data. The same sample size of 100k was used in many inconsistency simulation studies: [Aguaron and Moreno-Jimenez \[8\]](#), [Peláez and Lamata \[94\]](#), [Alonso and Lamata \[9\]](#), and [Vargas \[135\]](#).

- 500k sample: This is a very large sample the statistics of which will be used as final verification for both the 10k sample statistics and fit as well as the 100k sample statistics. It is interesting to note that in addition to the 100k sample, [Alonso and Lamata \[9\]](#) also generated a 500k sample and found no significant differences between the data from the tow sample sets.

## 4.2 Interpretation of the K-S and other GOF Tests

There are a few concepts to keep in mind while reviewing the distribution fit results.

The K-S test measures the maximum absolute deviation between the empirical distribution (distribution of simulation results) and a “standard theoretical” distribution such as the normal, lognormal, beta, or gamma distributions. Therefore, the larger the sample size, regardless of the GOF, the larger is the cumulative deviation and the worse is the K-S fit result. Accordingly, the smaller the K-S statistic of a fit is, the better the fit results are and vice versa. This phenomenon is highlighted by [Law and Kelton \[105\]](#) who say “This is an unfortunate property of these tests”.

Distribution fitting is done through hypothesis testing where the null hypothesis  $H_0$  is tested against the alternative hypothesis  $H_A$  as follows:

- $H_0$ : The empirical distribution and the “standard” distribution are the same.
- $H_A$ : The empirical distribution and the “standard” distribution are different.

If the fit is good, the recommendation with regard to the null hypothesis is “Do Not Reject”. If the fit is poor or not good enough at a particular significance level, the recommendation with regard to the null hypothesis is “Reject”. Once the distribution fitting is performed, a GOF statistic is calculated. Next, the hypothesis testing is done at a particular significance level  $\alpha$  and a corresponding critical value from the GOF statistic table is obtained. If the GOF statistic is greater than the critical value, the recommendation would be to reject the null hypothesis. If the GOF statistic is smaller than the critical value, the recommendation would be to not reject the null hypothesis.

In the case of this research, the desired outcome with regard to distribution fitting, is the null hypothesis (a recommendation of “Do Not Reject”). However, the significance level  $\alpha$  is the probability of rejecting the null hypothesis when it is true.

Therefore, higher alpha levels correspond to more stringent requirements of not rejecting the null hypothesis and accordingly higher significance in the GOF. This is the reverse of a situation when the alternative hypothesis (not the null) is the research hypothesis and the desired outcome. In that case lower alpha levels would imply higher significance.

In summary, the facts to keep in mind while reviewing the distribution fit results are:

- The smaller the K-S statistic, the higher is the GOF
- The larger the sample, the more difficult it is to obtain a reasonable GOF statistic
- The minimum sample corresponding to 99% precision and 99% confidence required for the K-S GOF test as indicated by [Gibbons \[104\]](#) is 6642
- A recommendation of “Do Not Reject” with regard to the null hypothesis is the desired outcome for distribution fitting
- The higher the alpha level for the “Do Not Reject” recommendation, the better the fit is which corresponds to higher confidence in the GOF

All of the distribution fitting, the fit results including fit parameters, CDF and quantile functions as well as goodness of fit test statistics were obtained using the distribution fitting software package EasyFit Professional Version 5.5.



### 4.3 Summary of Results

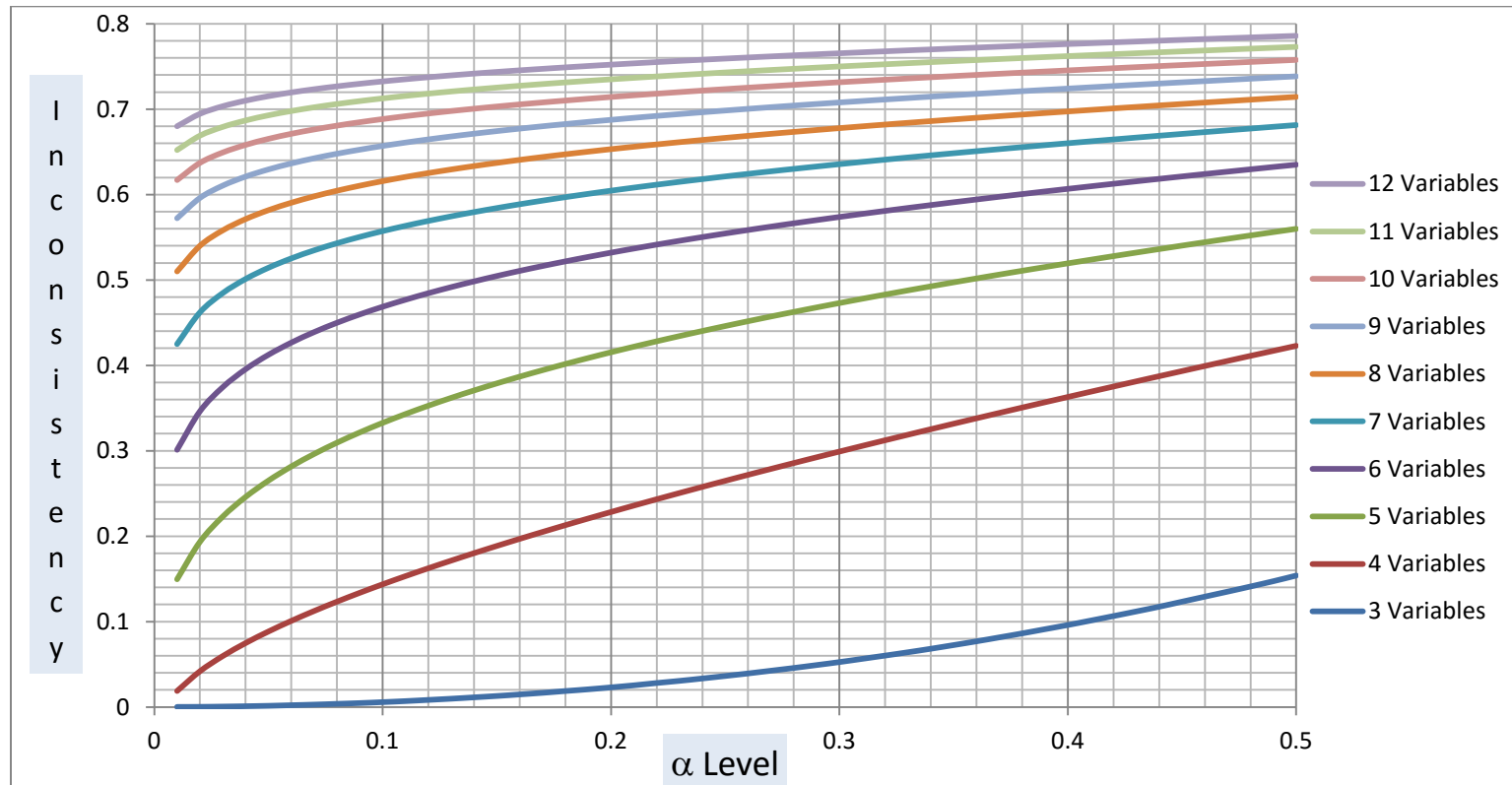
Following is a summary of the highlights of the research results.

- The inconsistency thresholds were defined for  $n = 3 - 12$  and corresponding fitted distributions were obtained.
- For each of the fitted distributions, the equations for the cumulative distribution and the quantile functions along with their specific set of parameters were identified.
- For  $n = 3$ , the fitted distribution is 3-parameter generalized gamma.
- For  $n = 4 - 12$ , the fitted distribution is Johnson SB.
- The GOF results are superb:
  - For  $n = 3$ , the GOF is “Do Not Reject” at all significance ( $\alpha$ ) levels (0.01, 0.02, .05, 0.1, and 0.2) for the K-S GOF test.
  - For  $n = 4 - 12$ , the GOF is “Do Not Reject” at all significance ( $\alpha$ ) levels (0.01, 0.02, .05, 0.1, and 0.2) for all GOF tests (K-S, Anderson-Darling, and Chi-Squared)
- All data verifications were performed with satisfactory outcome:
  - No significant difference was found between the 10k and 100k simulation data.
  - No significant difference was found between the 100k and 500k simulation data.

- No significant difference was found between the results from the fit equations and 500k simulation data.

RSV is based on the sum of the variances of the decision variable. The correlation among the pairwise comparisons variables decreases as the number of variables increases. This is accompanied by a decrease in variance as well. Consequently, when the number of variables reaches 13, the required growth in variance is no longer sufficient to provide the necessary increase for a new set of RSV values suitable for the new level (13). Therefore, the RSV measure cannot be used for calculations involving variables higher than 12. For detailed explanation of this phenomenon, refer to section 4.5.2.

Figure 4 below shows the inconsistency threshold limits for  $n = 3 - 12$  variables at  $\alpha$  from 0.01 to 0.5.



**Figure 4: Inconsistency Threshold Limits for 3 – 12 Decision Variables**

Table 4.1 below gives the inconsistency threshold limits for 3 – 12 variables at  $\alpha$  from 0.01 to 0.25 in 0.01 intervals.

Percentile ( $\alpha$ )	Number of variables									
	3	4	5	6	7	8	9	10	11	12
0.01	0.0001	0.0188	0.1495	0.3012	0.4249	0.5100	0.5723	0.6170	0.6521	0.6800
0.02	0.0002	0.0418	0.1934	0.3462	0.4620	0.5400	0.5961	0.6371	0.6690	0.6945
0.03	0.0005	0.0596	0.2230	0.3745	0.4847	0.5581	0.6105	0.6493	0.6793	0.7034
0.04	0.0009	0.0748	0.2460	0.3955	0.5012	0.5713	0.6211	0.6582	0.6868	0.7100
0.05	0.0014	0.0884	0.2651	0.4124	0.5143	0.5818	0.6295	0.6653	0.6928	0.7152
0.06	0.0021	0.1008	0.2816	0.4266	0.5253	0.5904	0.6365	0.6712	0.6978	0.7196
0.07	0.0028	0.1124	0.2963	0.4390	0.5347	0.5979	0.6425	0.6763	0.7022	0.7234
0.08	0.0037	0.1233	0.3095	0.4499	0.5430	0.6045	0.6478	0.6807	0.7060	0.7267
0.09	0.0046	0.1337	0.3215	0.4597	0.5505	0.6104	0.6526	0.6848	0.7095	0.7298
0.10	0.0057	0.1437	0.3327	0.4686	0.5572	0.6157	0.6569	0.6884	0.7126	0.7325
0.11	0.0069	0.1532	0.3430	0.4769	0.5634	0.6206	0.6609	0.6918	0.7155	0.7350
0.12	0.0083	0.1625	0.3528	0.4845	0.5691	0.6252	0.6646	0.6949	0.7182	0.7374
0.13	0.0097	0.1714	0.3620	0.4916	0.5745	0.6294	0.6681	0.6978	0.7207	0.7396
0.14	0.0113	0.1801	0.3706	0.4983	0.5795	0.6334	0.6713	0.7005	0.7231	0.7417
0.15	0.0129	0.1886	0.3789	0.5047	0.5842	0.6371	0.6743	0.7031	0.7253	0.7436
0.16	0.0147	0.1969	0.3868	0.5106	0.5887	0.6406	0.6772	0.7056	0.7274	0.7455
0.17	0.0166	0.2050	0.3944	0.5163	0.5929	0.6440	0.6800	0.7079	0.7294	0.7473
0.18	0.0187	0.2129	0.4016	0.5218	0.5970	0.6472	0.6826	0.7101	0.7313	0.7490
0.19	0.0207	0.2207	0.4086	0.5270	0.6009	0.6502	0.6851	0.7122	0.7332	0.7506
0.20	0.0230	0.2283	0.4154	0.5320	0.6046	0.6532	0.6875	0.7143	0.7350	0.7521
0.21	0.0254	0.2359	0.4219	0.5368	0.6081	0.6560	0.6899	0.7163	0.7367	0.7537
0.22	0.0281	0.2433	0.4282	0.5414	0.6116	0.6587	0.6921	0.7182	0.7383	0.7551
0.23	0.0306	0.2506	0.4343	0.5459	0.6149	0.6614	0.6943	0.7200	0.7399	0.7565
0.24	0.0334	0.2578	0.4403	0.5502	0.6181	0.6639	0.6964	0.7218	0.7415	0.7579
0.25	0.0363	0.2648	0.4461	0.5544	0.6212	0.6664	0.6984	0.7235	0.7430	0.7592

**Table 4.1: Inconsistency Threshold Limits for 3 – 12 Decision Variables at  $\alpha = 0.01$  to  $\alpha = 0.25$**

Table 4.2 below gives the inconsistency threshold limits for 3 – 12 variables at  $\alpha$  from 0.26 to 0.50 in 0.01 intervals.

Percentile ( $\alpha$ )	Number of variables									
	3	4	5	6	7	8	9	10	11	12
<b>0.26</b>	0.0393	0.2719	0.4517	0.5585	0.6243	0.6688	0.7004	0.7252	0.7445	0.7605
<b>0.27</b>	0.0425	0.2788	0.4572	0.5625	0.6272	0.6711	0.7023	0.7268	0.7459	0.7618
<b>0.28</b>	0.0457	0.2856	0.4626	0.5664	0.6301	0.6734	0.7042	0.7284	0.7473	0.7630
<b>0.29</b>	0.0491	0.2924	0.4679	0.5701	0.6329	0.6756	0.7061	0.7300	0.7487	0.7643
<b>0.30</b>	0.0526	0.2991	0.4730	0.5738	0.6356	0.6778	0.7079	0.7315	0.7500	0.7654
<b>0.31</b>	0.0564	0.3057	0.4780	0.5774	0.6383	0.6799	0.7096	0.7330	0.7513	0.7666
<b>0.32</b>	0.0602	0.3123	0.4830	0.5809	0.6409	0.6820	0.7113	0.7345	0.7526	0.7677
<b>0.33</b>	0.0641	0.3188	0.4878	0.5843	0.6434	0.6840	0.7130	0.7359	0.7538	0.7689
<b>0.34</b>	0.0682	0.3253	0.4926	0.5877	0.6460	0.6860	0.7147	0.7374	0.7551	0.7700
<b>0.35</b>	0.0725	0.3317	0.4972	0.5910	0.6484	0.6880	0.7163	0.7388	0.7563	0.7710
<b>0.36</b>	0.0769	0.3380	0.5018	0.5942	0.6508	0.6899	0.7179	0.7401	0.7575	0.7721
<b>0.37</b>	0.0815	0.3443	0.5063	0.5974	0.6532	0.6918	0.7195	0.7415	0.7587	0.7732
<b>0.38</b>	0.0862	0.3506	0.5108	0.6006	0.6555	0.6937	0.7211	0.7428	0.7599	0.7742
<b>0.39</b>	0.0911	0.3568	0.5152	0.6037	0.6579	0.6955	0.7226	0.7441	0.7610	0.7752
<b>0.40</b>	0.0960	0.3630	0.5195	0.6067	0.6601	0.6973	0.7241	0.7454	0.7621	0.7762
<b>0.41</b>	0.1012	0.3691	0.5237	0.6097	0.6624	0.6991	0.7256	0.7467	0.7633	0.7772
<b>0.42</b>	0.1065	0.3752	0.5279	0.6126	0.6646	0.7009	0.7271	0.7480	0.7644	0.7782
<b>0.43</b>	0.1119	0.3813	0.5321	0.6155	0.6668	0.7026	0.7286	0.7492	0.7655	0.7792
<b>0.44</b>	0.1174	0.3873	0.5362	0.6184	0.6689	0.7044	0.7300	0.7505	0.7666	0.7802
<b>0.45</b>	0.1233	0.3933	0.5402	0.6213	0.6710	0.7061	0.7314	0.7517	0.7677	0.7812
<b>0.46</b>	0.1292	0.3993	0.5443	0.6241	0.6732	0.7078	0.7329	0.7529	0.7687	0.7821
<b>0.47</b>	0.1351	0.4052	0.5482	0.6268	0.6752	0.7094	0.7343	0.7541	0.7698	0.7831
<b>0.48</b>	0.1411	0.4112	0.5521	0.6296	0.6773	0.7111	0.7357	0.7553	0.7709	0.7840
<b>0.49</b>	0.1473	0.4170	0.5560	0.6323	0.6794	0.7128	0.7371	0.7565	0.7719	0.7849
<b>0.50</b>	0.1539	0.4229	0.5599	0.6350	0.6814	0.7144	0.7384	0.7577	0.7729	0.7859

**Table 4.2: Inconsistency Threshold Limits for 3 – 12 Decision Variables at  $\alpha = 0.26$  to  $\alpha = 0.50$**

The goodness of fit test results for the RSV measure for inconsistency using the Generalized Gamma distribution at  $n = 3$ , and Johnson SB distribution at  $n = 4 - 12$  are summarized in table 4.3 below.

<b>n</b>	<b>Sample Size</b>	<b>Distribution</b>	<b>KS statistic</b>	<b>Level of Confidence for GOF Test</b>
<b>3</b>	10k	Gen. Gamma	0.00540	Do not reject at all significance levels - KS GOF test
<b>4</b>	10k	Johnson SB	0.00752	Do not reject at all significance levels - KS GOF test
<b>5</b>	10k	Johnson SB	0.00884	Do not reject at all significance levels - All GOF tests
<b>6</b>	10k	Johnson SB	0.00577	Do not reject at all significance levels - All GOF tests
<b>7</b>	10k	Johnson SB	0.00446	Do not reject at all significance levels - All GOF tests
<b>8</b>	10k	Johnson SB	0.00448	Do not reject at all significance levels - All GOF tests
<b>9</b>	10k	Johnson SB	0.00519	Do not reject at all significance levels - All GOF tests
<b>10</b>	10k	Johnson SB	0.00541	Do not reject at all significance levels - All GOF tests
<b>11</b>	10k	Johnson SB	0.00586	Do not reject at all significance levels - All GOF tests
<b>12</b>	10k	Johnson SB	0.00490	Do not reject at all significance levels - All GOF tests

**Table 4.3: Goodness of Fit for Variables 3-12**

The threshold limits shown in figure 4 and tables 4.1 and 4.2 as well as the goodness of fit results summarized in table 4.3 are based on 10,000 simulation runs of randomly generated pairwise comparison matrices. However, tests were conducted for 100,000 and 500,000 simulation runs and no significant deviations were observed in the results.

Table 4.4 below summarizes the results of those data verification runs.

<b>n</b>	<b>10k &amp; 500k</b>	<b>100k &amp; 500k</b>	<b>100k &amp; Fit</b>	<b>500k &amp; Fit</b>
<b>3</b>	0.0009	0.0002	0.0009	0.0005
<b>4</b>	0.0012	0.0007	0.0007	0.0032
<b>5</b>	0.0015	0.0005	0.0005	0.0018
<b>6</b>	0.0011	0.0011	0.0011	0.0010
<b>7</b>	0.0007	0.0004	0.0004	0.0006
<b>8</b>	0.0010	0.0004	0.0004	0.0002
<b>9</b>	0.0009	0.0002	0.0002	0.0002
<b>10</b>	0.0007	0.0003	0.0003	0.0001
<b>11</b>	0.0005	0.0003	0.0003	0.0002
<b>12</b>	0.0003	0.0003	0.0003	0.0001

**Table 4.4: Data Verification – Average Absolute Deviation in First 50 Percentiles**

The purpose of data verification is twofold:

1. Build confidence in data precision, and
2. Check the validity of the CDF and its inverse derived from the distribution fitting

**Data Precision:** Statistical rules generally indicate that the precision of a variable's estimation proportionally increases with sample size [\[136\]](#). Assuming the 500k sample to be the most precise, results of the other sample sizes (10k and 100k) were

compared to it. As shown in the first two columns of table 4.4, no significant differences were found.

**Validity of Fit Equations:** Data generated from the CDF and its inverse, which were both derived from the fitted distributions, were compared to both the 100k and 500k sample data. As shown in the last two columns of table 4.4, no significant differences were found.



## 4.4 Fitted Distributions

### Generalized Gamma

Generalized gamma is the fitted distribution for  $n = 3$ . The distribution has three parameters;  $k$ ,  $\alpha$ , and  $\beta$ . Those values for the best fit for  $n = 3$  are given in table 4.5.

Generalized Gamma does not have a closed form for the Cumulative Distribution Function (CDF). Therefore, the percentile of the inconsistency in quantified judgments for  $n = 3$  or the threshold limits for a given percentile cannot be obtained analytically. Table 4.1 and 4.2 can be used for that purpose.

### Johnson SB

The fitted distribution for  $n = 4 - 12$  is Johnson SB. It has four parameters;  $\gamma$ ,  $\delta$ ,  $\lambda$ , and  $\xi$ . The parameter values for the best fit for  $n = 4 - 12$  are given in table 4.5. Those values can be used for calculation of the percentile of the inconsistency (the  $\alpha$  level) in quantified judgments by using the Johnson SB Cumulative Distribution Function (CDF). The values can also be used for calculation of the inconsistency threshold at the desired  $\alpha$  level by using the Johnson SB Quantile (inverse CDF) function. The calculated threshold levels are also listed for  $\alpha = 0.01$  to  $0.5$  in tables 4.1 and 4.2.

### Cumulative distribution function for Johnson SB

$$F(x) = \Phi\left(\gamma + \delta \ln \frac{z}{1-z}\right) \quad (4.1)$$

where:

$$z = \frac{x - \xi}{\lambda} \quad (4.2)$$

$\Phi(\cdot)$  is the cumulative distribution function for standard normal distribution

$x$  is inconsistency calculated as RSV in the actual data obtained from quantified judgments

$$RSV = \sqrt{\sum_{i=1}^n \sigma_i^2} \quad (4.3)$$

### Quantile (Inverse cumulative distribution) function for Johnson SB

$$F^{-1}(P) = \frac{\lambda \exp\left(\frac{\Phi^{-1}(P) - \gamma}{\delta}\right)}{1 + \exp\left(\frac{\Phi^{-1}(P) - \gamma}{\delta}\right)} + \xi \quad (4.4)$$

where  $\Phi^{-1}(\cdot)$  is the inverse cumulative distribution function for standard normal distribution.

Generalized gamma is the fitted distribution for  $n = 3$ . The distribution has three parameters;  $k$ ,  $\alpha$ , and  $\beta$ . Those values for the best fit for  $n = 3$  are given in table 4.5 below. The fitted distribution for  $n = 4 - 12$  is Johnson SB. It has four parameters;  $\gamma$ ,  $\delta$ ,  $\lambda$ , and  $\xi$ . The parameter values for the best fit for  $n = 4 - 12$  are given in table 4.5 below.

		Distribution Parameters			
<b>n</b>	<b>Fitted Distribution</b>	<b><math>k</math></b>	<b><math>\alpha</math></b>	<b><math>\beta</math></b>	
<b>3</b>	Generalized Gamma	10.691	0.04501	0.70493	
		<b><math>\gamma</math></b>	<b><math>\delta</math></b>	<b><math>\lambda</math></b>	<b><math>\xi</math></b>
<b>4</b>	Johnson SB	-	-	-	-
		0.24235	0.87961	0.83692	0.05284
<b>5</b>	Johnson SB	-1.1717	1.2564	0.94894	-
					0.12107
<b>6</b>	Johnson SB	-2.3764	1.7333	1.1499	-
					0.28209
<b>7</b>	Johnson SB	-4.3678	2.4409	1.6113	-
					0.69925
<b>8</b>	Johnson SB	-7.2967	3.1490	2.5136	-1.5737
<b>9</b>	Johnson SB	-7.5056	3.4448	2.0710	-1.1220
<b>10</b>	Johnson SB	-13.443	4.3652	4.9027	-3.9295
<b>11</b>	Johnson SB	-9.1512	4.1515	1.9309	-
					0.96609
<b>12</b>	Johnson SB	-8.6057	4.2984	1.5263	-
					0.55881

**Table 4.5: Fitted Distribution Parameters**

#### 4.4.1 Using the Johnson SB CDF and Quantile Functions

The  $\alpha$  level is determined for a calculated inconsistency by using the Johnson SB CDF as follows:

##### Input for CDF:

- The inconsistency value ( $x$ )
- The Johnson SB parameters corresponding to the number of variables ( $\gamma, \delta, \lambda, \xi$ )

##### Output of CDF:

- The  $\alpha$  level corresponding to a given inconsistency

Calculations can be performed in Excel as follows:

$$\alpha = \text{NORMSDIST}(\gamma + \delta * \text{LN}(Z/(1-Z))) \quad (4.5)$$

where:

NORMSDIST is an Excel function that returns the standard normal cumulative distribution function

$\gamma$  is the gamma parameter for Johnson SB distribution

$\delta$  is the delta parameter for Johnson SB distribution

$$z = \frac{x - \xi}{\lambda} \quad (4.6)$$

where  $x$  is the inconsistency value for which  $\alpha$  is to be calculated

$$x = \sqrt{\sum_{i=1}^n \sigma_i^2} \quad (4.7)$$

$\lambda$  is the lambda parameter for Johnson SB distribution

$\xi$  is the xi parameter for Johnson SB distribution

Using table 4.5, for  $n = 5$ :  $\gamma = -1.1717$ ,  $\delta = 1.2564$ ,  $\lambda = 0.94894$ ,  $\xi = -0.121407$ , and

choose  $x = 0.32$

Substituting in 4.6 above:

$$z = \frac{x - \xi}{\lambda} = \frac{0.32 - (-0.121407)}{0.94894} = 0.465158 \quad (4.8)$$

Substituting in 4.5 above:

$$\alpha = \text{NORMSDIST}(-1.1717 + 1.12564 * \text{LN}(0.465158 / (1 - 0.465158))) = 0.092$$

This is a reasonable  $\alpha$  value for  $n = 5$ , and therefore this level of inconsistency ( $x = 0.32$ )

could be accepted.

The inconsistency threshold is determined for a given  $\alpha$  level by using the Johnson SB Quantile (Inverse CDF) function as follows:

**Input for Quantile (Inverse CDF) function:**

- $\alpha$  level
- The Johnson SB parameters corresponding to the number of variables

**Output of Quantile function:**

- Inconsistency threshold corresponding to the  $\alpha$  level

Calculations can be performed in Excel as follows:

$$\text{Inconsistency} = \lambda * \text{EXP}((\text{NORMSINV}(\alpha) - \gamma) / \delta) / (1 + \text{EXP}((\text{NORMSINV}(\alpha) - \gamma) / \delta)) + \xi \quad (4.9)$$

where:

NORMSDIST is an Excel function that returns the inverse of the standard normal cumulative distribution.

$\gamma$  is the gamma parameter for Johnson SB distribution

$\delta$  is the delta parameter for Johnson SB distribution

$\lambda$  is the lambda parameter for Johnson SB distribution

$\xi$  is the xi parameter for Johnson SB distribution

$\alpha$  is the  $\alpha$  level for which the inconsistency threshold is to be calculated

Using table 4.5, for  $n = 7$ :  $\gamma = -4.3678$ ,  $\delta = 2.4409$ ,  $\lambda = 1.6113$ ,  $\xi = -0.69925$ , and choose

$\alpha = 0.15$

$$\text{Inconsistency} = \lambda * \text{EXP}((\text{NORMSINV}(\alpha) - \gamma) / \delta) / (1 + \text{EXP}((\text{NORMSINV}(\alpha) - \gamma) / \delta)) + \xi$$

$$\text{Inconsistency} = 1.6113 * \text{EXP}((\text{NORMSINV}(0.15) - (-4.3678)) /$$

$$2.4409) / (1 + \text{EXP}((\text{NORMSINV}(0.15) - (-4.3678)) / 2.4409)) + (-0.69925) = 0.587 \text{ which is}$$

maximum inconsistency to achieve  $\alpha = 0.15$  for  $n = 7$ .

## 4.5 Discussion of Results

### 4.5.1 Note on the Distribution for $n = 3$

For  $n = 3$ , the fitted distribution is Generalized Gamma. This is an exception since for the entire range beyond  $n = 3$  ( $n = 4 - 12$ ), a single distribution (Johnson SB) was found to have the best fit. The  $n = 3$  exception is not surprising. This phenomenon (and same curve shape) was reported by previous simulation studies related to inconsistency measures. Relevant examples include [Lane and Verdini \[72\]](#) who generated the entire null distribution for  $n = 3$  for Saaty's inconsistency measure. Almost identical results were obtained by [Bozoki and Rapcsak \[13\]](#) who also generated a huge sample for  $n = 3$ . [Kretchik and Ra \[86\]](#) obtained similar results for their sum of the inverse column sums consistency measure. Even though those tests were conducted on a different inconsistency measure, it is reassuring that the distribution of inconsistencies in human judgments was found to have the same shape as the distributions we identified in this research.



#### 4.5.2 Change of RSV's Behavior at $n = 13$

By design (see section 3.1), the RSV parameter was developed to be a function of the number of decision variables. This implies that RSV should comply with the following condition:

$$RSV_{\alpha}(n + 1) > RSV_{\alpha}(n) \quad \forall \alpha = 0.01 - 1 \quad (4.10)$$

This means that RSV magnitudes corresponding to a number of decision variables should be higher in value than those corresponding to the next lower number of decision variables at any given  $\alpha$ . RSV complies with the above condition for the range  $n = 3 - 12$ . However, at  $n = 13$ , RSV values become smaller than those for  $n = 12$ . This change in expected behavior is explained below.

Pairwise comparison decision variables such as those used in AHP/HDM have the following mathematical/statistical properties:

1. They are correlated to each other, directly and indirectly, through pairwise comparisons, the values of which are used to calculate the variables' relative weights
2. Each value of each variable is given only relative to those of others and not as an absolute value.
3. The variables' values are usually represented as parts of one or percentages of a whole.

The RSV consistency measure is based on the sum of the variances of the decision variables represented by the columns of the orientation table. Because the variables are correlated, the variance of their sum is the sum of their covariances:

$$Var\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n Var(X_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n Cov(X_i, X_j) \quad (4.11)$$

Because of the pairwise data representation form and the properties mentioned above, equation (4.11) always yields zero. The two terms of the equation are directly proportional to each other: equal in magnitude and opposite in sign with the first term always being positive. The columns of the orientation table produce the first (positive term) and the rows produce the covariance term. Therefore, if the above equation is applied to both the columns and rows of the orientation table, the total will add to zero.

It is important to note that the covariance (2nd term) reflects the correlations among the variables, and therefore is directly affected by the strength of such correlations. In the case of two pairwise comparison variables, the relationship is deterministic since knowing the value of one enables the determination of the second. Not only is the correlation in this case definitive but is also negative, reaching the maximum correlation coefficient of -1, since an increase in either would necessitate a decrease in the other. The relationship among three variables may not be as direct as in the case of two, but at least two of the three variables would be strongly correlated. By extending this

argument, the conclusion becomes that correlation increases with fewer variables and decreases with larger number of variables.

RSV is computed using the first (variance) term of the equation only, but is heavily affected by the second (covariance) term due to the fixed sum (zero) nature of the equation. The required proportional growth in RSV as the number of variables increases comes from the increase in the variance term brought about by the contribution of an additional variable (additional column in the orientation table). Nonetheless, as the number of variables increases, the correlation among them decreases which contributes to the decrease of the covariance term, bringing about a proportional decrease in the variance term as well. This continues to happen until the growth in RSV brought about by the contribution of an additional variable is no longer sufficient to compensate for the decrease brought about the shrinking covariance. At that point, which occurs at 13 variables, it becomes infeasible to use RSV as a consistency measure.

To validate the above explanation of RSV's change in behavior, a simulation study was carried out to characterize the variance. For  $n = 3 - 13$ , 100,000 each randomly filled matrices were used to calculate the average variance.

The results show that for  $n = 3 - 6$ , the average variance increases at a diminishing rate and reaches its maximum at  $n = 6$ . The increase which is augmented by the

additional contribution of the variance of an extra element account for the decreasing but substantial  $\Delta y$  between the successive curves of  $n = 3, 4, 5, \text{ and } 6$  as Figure 4 shows.

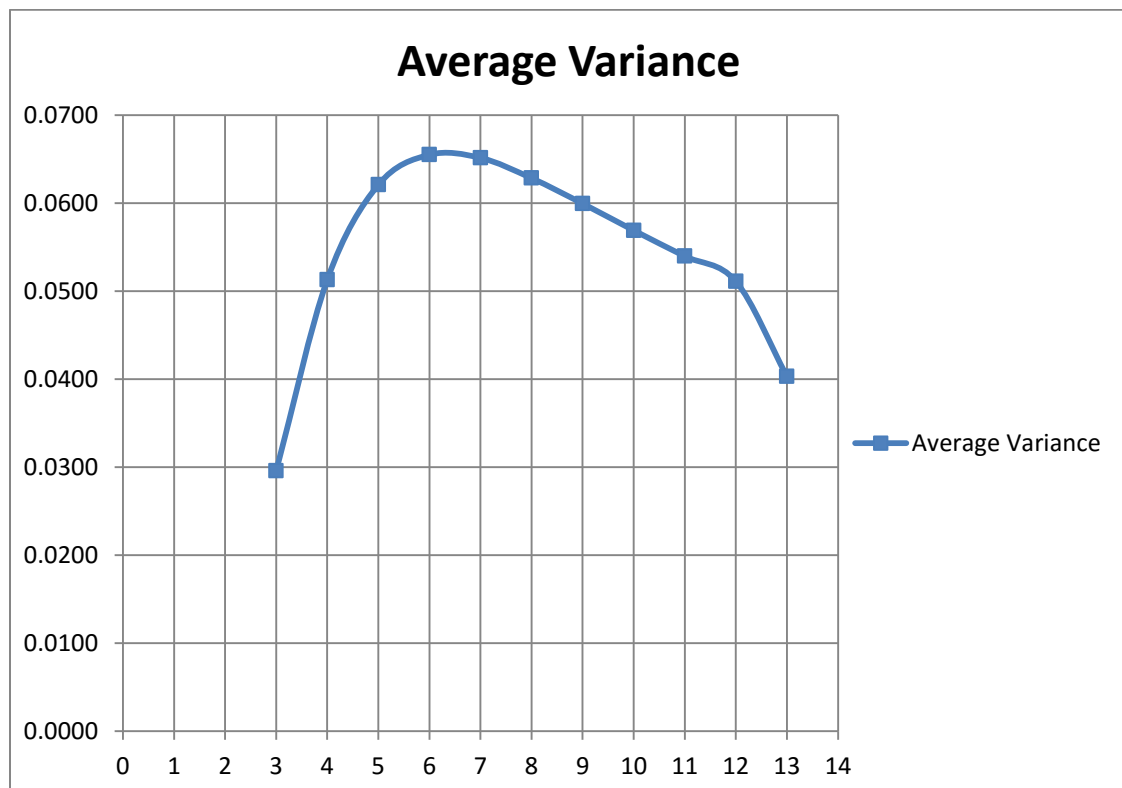
For  $n = 7 - 13$ , the average variance steadily decreases by about 5%. At  $n = 13$ , the decrease is 21%. The steady decrease in this range is offset by the additional contribution of the variance of an extra element. However,  $\Delta y$  between successive curves of elements in this range continues to decrease as Figure 4 shows.

At  $n = 13$ , the 21% drop in average variance causes the sum of the variances to become 85% of that at  $n = 12$  and the necessary increase in the parameter for the new level is lost.

The results of the simulation study, shown on the next two pages, validate the assumptions and confirm the explanation given in this section.

<i>n</i>	Average variance	% of previous rank	Total variance	% of previous rank
3	0.0296		0.0887	
4	0.0513	173%	0.2052	231%
5	0.0621	121%	0.3104	151%
6	0.0655	105%	0.3930	127%
7	0.0651	99%	0.4560	116%
8	0.0628	96%	0.5027	110%
9	0.0600	95%	0.5396	107%
10	0.0569	95%	0.5690	105%
11	0.0540	95%	0.5937	104%
12	0.0511	95%	0.6134	103%
13	0.0403	79%	0.5241	85%

**Table 4.6: Average Variance and Related Calculations for  $n = 3 - 13$**



**Figure 5: Average Variance for  $n = 3 - 13$**

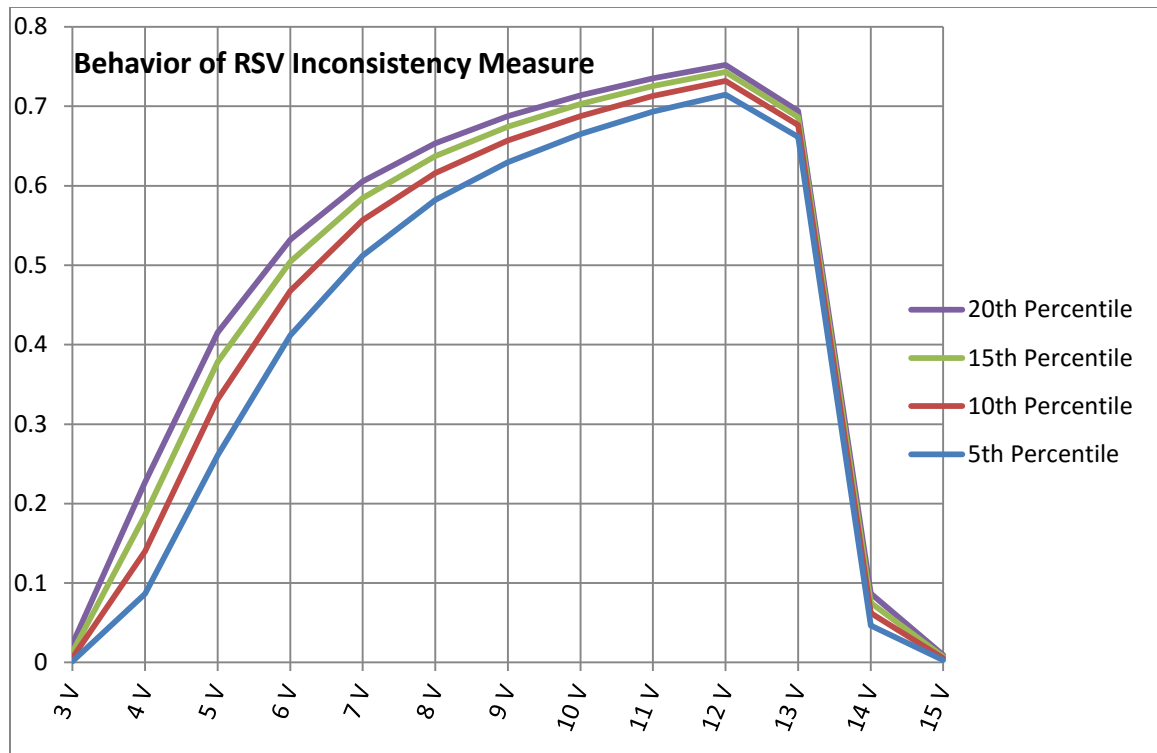


Figure 6: RSV Behavior for  $n = 3 - 13$

## 4.6 Assessment of Current Consistency Measure

The current consistency measure in HDM is based on the arithmetic mean of the standard deviation of the decision variables:

$$Inconsistency = \frac{1}{n} \sum_{i=1}^n \sigma_i \quad (4.12)$$

A comparative study was carried out to assess the current consistency measure in light of the proposed measure. Data sets of typical values were used to produce inconsistency equal to the 10% limit of the current measure. RSV, the proposed measure, was then used to calculate the inconsistency for the same data sets. This was done for the range of decision elements 3—12. The results are as follows:

- For  $n = 3$ , the 10% fixed limit corresponds to  $\alpha = 0.24$
- For  $n = 4$ , the 10% fixed limit corresponds to  $\alpha = 0.03$
- For  $n \geq 5$ , the 10% fixed limit corresponds to  $\alpha < 0.01$

Based on the above, it is concluded that the recommended fixed limit of 10% for the current inconsistency measure is very conservative for  $n > 3$ , and it becomes increasingly so as  $n$  becomes larger.

The orientation table of the HDM numerical example given in section 1.4.3 is reintroduced here to allow for comparison of the values of the proposed RSV and those of the current measure (4.12) above.

Recall, RSV is the square root of the sum of variances while the current measure is the average standard deviation.

#	Orientation	Before normalization					Normalized			
		A	B	C	D	Sum	A	B	C	D
1	ABCD	1.81	3.31	0.74	1.00	6.85	0.26	0.48	0.11	0.15
2	ABDC	3.61	6.62	1.00	3.11	14.35	0.25	0.46	0.07	0.22
3	ACBD	1.88	2.13	0.60	1.00	5.61	0.34	0.38	0.11	0.18
4	ACDB	1.56	1.00	0.50	0.68	3.73	0.42	0.27	0.13	0.18
5	ADBC	3.50	4.48	1.00	3.03	12.00	0.29	0.37	0.08	0.25
6	ADCB	1.02	1.00	0.28	0.88	3.19	0.32	0.31	0.09	0.28
7	BACD	2.30	5.65	0.74	1.00	9.69	0.24	0.58	0.08	0.10
8	BADC	3.60	8.84	1.00	3.11	16.56	0.22	0.53	0.06	0.19
9	BCAD	1.16	3.86	0.86	1.00	6.88	0.17	0.56	0.13	0.15
10	BCDA	1.00	6.58	1.47	1.99	11.04	0.09	0.60	0.13	0.18
11	BDCA	1.00	4.94	0.75	2.32	9.01	0.11	0.55	0.08	0.26
12	BDAC	3.11	13.17	1.00	6.19	23.47	0.13	0.56	0.04	0.26
13	CABD	1.16	2.13	0.87	1.00	5.16	0.23	0.41	0.17	0.19
14	CADB	0.78	1.00	0.58	0.68	3.04	0.26	0.33	0.19	0.22
15	CBAD	1.16	2.84	0.81	1.00	5.80	0.20	0.49	0.14	0.17
16	CBDA	1.00	4.23	1.20	1.99	8.42	0.12	0.50	0.14	0.24
17	CDAB	0.55	1.00	0.80	1.09	3.43	0.16	0.29	0.23	0.32
18	CDBA	1.00	2.46	1.23	1.66	6.34	0.16	0.39	0.19	0.26
19	DACB	0.88	1.00	0.28	1.76	3.92	0.23	0.25	0.07	0.45
20	DABC	2.44	4.48	1.00	4.86	12.78	0.19	0.35	0.08	0.38
21	DBAC	3.11	7.65	1.00	5.17	16.93	0.18	0.45	0.06	0.31
22	DBCA	1.00	3.34	0.75	2.26	7.35	0.14	0.45	0.10	0.31
23	DCAB	0.55	1.00	0.41	1.27	3.22	0.17	0.31	0.13	0.39
24	DCBA	1.00	2.46	0.70	2.17	6.32	0.16	0.39	0.11	0.34
Mean							0.21	0.43	0.11	0.25
$\sigma^2$							0.00591	0.0106	0.00218	0.00710

**Table 4.7: Orientation Table from HDM Numerical Example in Section 1.4.3**



## Inconsistency Calculations

### Proposed Measure:

$$RSV = \sqrt{\sum_1^n \sigma_n^2} \quad (4.13)$$

$$RSV = \sqrt{0.00591 + 0.0106 + 0.00218 + 0.00710}$$

$$RSV = 0.161$$

Using table 4.1 for  $n = 4$ , this RSV value corresponds to  $\alpha = 0.12$ . This value is somewhat high considering  $n = 4$ . Therefore, it is recommended for the DM to revise the pairwise comparison assignments.

### Current Measure:

$$Inconsistency = \frac{1}{n} \sum_{i=1}^n \sigma_i \quad (4.14)$$

$$Inconsistency = \frac{1}{4} (\sqrt{0.00591} + \sqrt{0.0106} + \sqrt{0.00218} + \sqrt{0.00710})$$

$$Inconsistency = \frac{1}{4} (0.0769 + 0.1029 + 0.0467 + 0.0843)$$

$$Inconsistency = 7.77\%$$

The calculated inconsistency is below the maximum fixed 10% threshold, and therefore pairwise comparison assignments should be accepted.

## 5 CONCLUSION

The importance of measuring and controlling consistency in any AHP/HDM application cannot be overemphasized. Nonetheless, any consistency measure without meaningful thresholds remains mainly abstract and offers limited practical benefit. HDM's proposed new inconsistency measure, RSV, along with the thresholds established as a result of this research fulfill all the requirements previously established for a robust, useful, and practical consistency measure. RSV and its thresholds are:

- A function of the number of decision variables
- Derived using statistical hypothesis testing
- Linked to any desired  $\alpha$  levels

The thresholds allow decision makers who provide data through pairwise comparisons as well as decision makers who use or apply the decisions based on those pairwise choices to assess the soundness and validity of their decisions. Moreover, the thresholds allow the various decision makers to select a particular level of  $\alpha$  which is appropriate to the specific circumstances of the decision problem.

HDM with its judgement quantification methodology, sensitivity analysis, and the proposed robust inconsistency measure with inconsistency thresholds is one of the most complete methods available for multi-level, multi-criteria decisions.

## 6 CONTRIBUTION

The contribution of this research is to fill an important research gap identified through the literature review by defining the acceptable limits of inconsistency for any number of decision elements from 3 to 12 at any given  $\alpha$  level in HDM calculations.

A byproduct of this research includes two fundamental improvements to HDM's judgement quantification method to enhance its speed and efficiency while maintaining a high degree of accuracy. This is done by the development of new computation algorithms that drastically reduce computational burden thereby greatly increasing the method's speed and consequently making it truly practical. Sections 3.2.5 – 3.2.9 contain detailed explanations of these improvements.

## 7 LIMITATIONS

1. The RSV inconsistency measure cannot be used for pairwise comparisons in HDM models where  $n > 12$ . Because of the limitation of the mathematical formula of variance, thresholds for variables above 12 cannot be calculated using the current method for RSV.

However, this is not a major limitation because:

- a. Research [\[137\]](#) shows that as  $n$  increases, the ability of a DM to improve judgment decreases, and therefore consistency deteriorates. [Saaty and Ozdemir \[137\]](#) argue that at higher  $n$ , consistency improvement requires small changes to which most human beings are insensitive. [Saaty and Ozdemir \[137\]](#) state “we conclude that to serve both consistency and redundancy, it is best to keep the number of elements seven or less”.
- b. There does not seem to be a practical need for pairwise comparisons of more than 12 elements. A review of all the HDM application articles discussed in section 2.2.4 shows that in all the models presented, the largest number of alternatives evaluated as a group is nine. Table 7.1 on page 130 lists the maximum number of alternatives forming a single group in every HDM application article.
- c. Pairwise comparisons of more than 12 elements are not considered realistic. Twelve elements represent 66 pairwise comparisons which is

already a formidable number of relationships to track. The number of comparisons jumps to 78 for 13 elements, 91 for 14 and 105 for 15.

It is very difficult if not impossible for a decision maker to evaluate more than 66 comparisons reliably.

- d. If there are more than 12 elements in the decision problem, the elements can be divided into groups and evaluated by chainwise comparisons [Ra \[138\]](#). This way each group can be limited to less than 12 elements and the consistency measure developed in this dissertation can be applied repeatedly.

- 2. The inconsistency measure developed in this dissertation is applicable to HDM as stated in the research objective. RSV is the square root of the sum of variances of the means of  $n$  decision variables obtained in  $n$  factorial orientations. This computation is used by the HDM judgment quantification method, not other methods.

Table 7.1 below lists the maximum number of alternatives forming a single group in every HDM application article.

#	HDM Application Article (Section 2.2.4)	Maximum Number of Alternatives Evaluated Together
1	<a href="#">Khan [55]</a>	9
2	<a href="#">Gerdsri [56]</a>	7
3	<a href="#">Chen, <i>et al.</i> [58]</a>	5
4	<a href="#">Gerdsri and Kocaoglu [59]</a>	7
5	<a href="#">Fenwick, <i>et al.</i> [60]</a>	7
6	<a href="#">Kodali, <i>et al.</i> [61]</a>	9
7	<a href="#">Cowan, <i>et al.</i> [62]</a>	7
8	<a href="#">Daim, <i>et al.</i> [63]</a>	7
9	<a href="#">Harell and Daim [64]</a>	5
10	<a href="#">Kennedy and Daim [65]</a>	5
11	<a href="#">Wang, <i>et al.</i> [66]</a>	6
12	<a href="#">Fenwick and Daim [67]</a>	4
13	<a href="#">Sheikh, <i>et al.</i> [68]</a>	8
14	<a href="#">Daim, <i>et al.</i> [69]</a>	5
15	<a href="#">Lee, <i>et al.</i> [70]</a>	4

**Table 7.1: List of Maximum Number of Alternatives Evaluated All Together in Each of the HDM Application Articles**

## 8 FUTURE WORK

The development of RSV as a new measure of inconsistency in HDM, with thresholds at desired  $\alpha$  levels, and defined as a function of the number of decision variables has met all the goals set out in the proposal of this dissertation. It has also identified the opportunity for future work to extend the inconsistency measure to one which:

- Can be used for calculations involving more than twelve decision variables
- Is independent of the judgment quantification method, thereby eliminating the need for the calculation of the variance of the means of  $n$  variables in  $n$  factorial orientations.
- Is universally applicable to any pairwise-comparison based method.

## REFERENCES

- [1] T. L. Saaty, "Decision-making with the AHP: Why is the principal eigenvector necessary," *European Journal of Operational Research*, vol. 145, pp. 85-91, 2003.
- [2] D. F. Kocaoglu, "A Participative Approach to Program Evaluation," *IEEE Transactions on Engineering Management*, vol. EM-30, pp. 112-118, August 1983.
- [3] T. L. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*: McGraw-Hill International Book Co., 1980.
- [4] D. I. Cleland and D. F. Kocaoglu, *Engineering Management*. New York: McGraw-Hill, 1981.
- [5] P. Gerdri and D. Kocaoglu, "A systematic approach to developing national technology policy and strategy for emerging technologies: A case study of nanotechnology for Thailand's agriculture industry," in *PICMET 2009 - 2009 Portland International Conference on Management of Engineering and Technology, August 2, 2009 - August 6, 2009*, Portland, OR, United states, 2009, pp. 447-461.
- [6] R. D. Holder, "Some Comments on the Analytic Hierarchy Process," *The Journal of the Operational Research Society*, vol. 41, pp. 1073-1076, 1990.
- [7] C. Murphy, "Limits on the analytic hierarchy process from its consistency index," *European Journal of Operational Research*, vol. 65, pp. 138-139, 1993.
- [8] J. Aguaron and J. M. Moreno-Jimenez, "The geometric consistency index: Approximated thresholds," *European Journal of Operational Research*, vol. 147, pp. 137-145, 2003.
- [9] J. A. Alonso and M. T. Lamata, "Consistency in the analytic hierarchy process: A new approach," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 14, pp. 445-459, 2006.
- [10] T. L. Saaty, "A scaling method for priorities in hierarchical structures," *Journal of Mathematical Psychology*, vol. 15, pp. 234-281, 1977.
- [11] E. H. Forman, "Random indices for incomplete pairwise comparison matrices," *European Journal of Operational Research*, vol. 48, pp. 153-155, 1990.
- [12] T. L. Saaty, *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*. Pittsburgh, PA: RWS Publications, 1994.
- [13] S. Bozoki and T. Rapcsak, "On Saaty's and Koczkodaj's inconsistencies of pairwise comparison matrices," *Journal of Global Optimization*, vol. 42, pp. 157-175, 2008.
- [14] B. Cavallo and L. D'Apuzzo, "A general unified framework for pairwise comparison matrices in multicriterial methods," *International Journal of Intelligent Systems*, vol. 24, pp. 377-398, 2009.
- [15] E. U. Choo and W. C. Wedley, "A common framework for deriving preference values from pairwise comparison matrices," *Computers and Operations Research*, vol. 31, pp. 893-908, 2004.



- [16] C.-C. Lin, "A revised framework for deriving preference values from pairwise comparison matrices," *European Journal of Operational Research*, vol. 176, pp. 1145-1150, 2007.
- [17] E. H. Forman and S. I. Gass, "The analytic hierarchy process-an exposition," *Operations Research*, vol. 49, pp. 469-86, 2001.
- [18] S. Opasanon and P. Lertsanti, "Impact analysis of logistics facility relocation using the analytic hierarchy process (AHP)," *International Transactions in Operational Research*, vol. 20, pp. 325-339, 2013.
- [19] S. Sipahi and M. Timor, "The analytic hierarchy process and analytic network process: an overview of applications," *Management Decision*, vol. 48, pp. 775-808, 2010.
- [20] Z. Xu, "A Practical Method for Improving Consistency of Judgement Matrix in the AHP," *Journal of Systems Science and Complexity* vol. 17, pp. 164-168, 2004.
- [21] H. Maleki and S. Zahir, "A Comprehensive Literature Review of the Rank Reversal Phenomenon in the Analytic Hierarchy Process," *Journal of Multi-Criteria Decision Analysis*, vol. 20, pp. 141-155, 2013.
- [22] C.-C. Lin, *et al.*, "Improving AHP for construction with an adaptive AHP approach (A3)," *Automation in Construction*, vol. 17, pp. 180-187, 2008.
- [23] R. Medjoudj, *et al.*, "Decision making on power customer satisfaction and enterprise profitability analysis using the Analytic Hierarchy Process," *International Journal of Production Research*, vol. 50, pp. 4793-4805, 2012.
- [24] N. Bhushan and K. Rai, *Strategic Decision Making: Applying the Analytic Hierarchy Process*: Springer, 2004.
- [25] A. Ishizaka, *et al.*, "AHPSort: an AHP-based method for sorting problems," *International Journal of Production Research*, vol. 50, pp. 4767-4784, 2012.
- [26] W. Ho, "Integrated analytic hierarchy process and its applications – A literature review," *European Journal of Operational Research*, vol. 186, pp. 211-228, 2008.
- [27] O. S. Vaidya and S. Kumar, "Analytic hierarchy process: An overview of applications," *European Journal of Operational Research*, vol. 169, pp. 1-29, 2006.
- [28] A. Ishizaka and A. Labib, "Review of the main developments in the analytic hierarchy process," *Expert Systems with Applications*, vol. 38, pp. 14336-14345, 2011.
- [29] T. L. Saaty, "Decision making with the analytic hierarchy process," *International Journal of Services Sciences* vol. 1, pp. 83 - 98, 2008.
- [30] D. Cao, *et al.*, "Modifying inconsistent comparison matrix in analytic hierarchy process: A heuristic approach," *Decision Support Systems*, vol. 44, pp. 944-953, 2008.
- [31] M. Brunelli, *et al.*, "A note on the proportionality between some consistency indices in the AHP," *Applied Mathematics and Computation*, vol. 219, pp. 7901-7906, 2013.

- [32] D. R. Anderson, *et al.*, *An Introduction to Management Science: Quantitative Approaches to Decision Making*: Cengage South-Western, 2010.
- [33] M. T. Lamata and J. I. Pelaez, "A method for improving the consistency of judgements," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, pp. 677-686, 2002.
- [34] G. Kou, *et al.*, "Enhancing data consistency in decision matrix: Adapting Hadamard model to mitigate judgment contradiction," *European Journal of Operational Research*, vol. 236, pp. 261-271, 2014.
- [35] S. Bortot and R. A. Marques Pereira, "Inconsistency and non-additive capacities: The Analytic Hierarchy Process in the framework of Choquet integration," *Fuzzy Sets and Systems*, vol. 213, pp. 6-26, 2013.
- [36] C. Lin, *et al.*, "A statistical approach to measure the consistency level of the pairwise comparison matrix," *Journal of the Operational Research Society*, 2013.
- [37] M. Brunelli and M. Fedrizzi, "Axiomatic properties of inconsistency indices for pairwise comparisons," *Journal of the Operational Research Society*, 2013.
- [38] D. Ergu, *et al.*, "A simple method to improve the consistency ratio of the pairwise comparison matrix in ANP," *European Journal of Operational Research*, vol. 213, pp. 246-259, 2011.
- [39] D. V. Budescu, *et al.*, "A Comparison of the Eigenvalue Method and The Geometric Mean Procedure for Ratio Scaling," *Applied Psychological Measurement*, vol. 10, pp. 69-78, 1986.
- [40] F. Zahedi, "A simulation study of estimation methods in the analytic hierarchy process," *Socio-Economic Planning Sciences*, vol. 20, pp. 347-354, 1986.
- [41] E. Noble and P. Sanchez, "A note on the information content of a consistent pairwise comparison judgment matrix of an AHP decision maker," *Theory and Decision*, vol. 34, pp. 99-108, 1993.
- [42] T. L. Saaty and L. G. Vargas, "Experiments on rank preservation and reversal in relative measurement," *Mathematical and Computer Modelling*, vol. 17, pp. 13-18, 1993.
- [43] C. Genest and L.-P. Rivest, "A Statistical Look at Saaty's Method of Estimating Pairwise Preferences Expressed on a Ratio Scale," *Journal of Mathematical Psychology*, vol. 38, pp. 477-496, 1994.
- [44] F. J. Carmone Jr, *et al.*, "A Monte Carlo investigation of incomplete pairwise comparison matrices in AHP," *European Journal of Operational Research*, vol. 102, pp. 538-553, 1997.
- [45] J. S. Finan and W. J. Hurley, "Analytic hierarchy process: Does adjusting a pairwise comparison matrix to improve the consistency ratio help?," *Computers and Operations Research*, vol. 24, pp. 749-755, 1997.
- [46] S. H. Zanakakis, *et al.*, "Multi-attribute decision making: A simulation comparison of select methods," *European Journal of Operational Research*, vol. 107, pp. 507-529, 1998.

- [47] R. Aull-Hyde, *et al.*, "An experiment on the consistency of aggregated comparison matrices in AHP," *European Journal of Operational Research*, vol. 171, pp. 290-295, 2006.
- [48] A. Ishizaka and M. Lusti, "How to derive priorities in AHP: a comparative study," *Central European Journal of Operations Research*, vol. 14, pp. 387-400, 2006.
- [49] N. J. Z. Mamat and J. K. Daniel, "Statistical analyses on time complexity and rank consistency between singular value decomposition and the duality approach in AHP: A case study of faculty member selection," *Mathematical and Computer Modelling*, vol. 46, pp. 1099-1106, 2007.
- [50] Y. Dong, *et al.*, "Consensus models for AHP group decision making under row geometric mean prioritization method," *Decision Support Systems*, vol. 49, pp. 281-289, 2010.
- [51] A. Ishizaka, *et al.*, "Influence of aggregation and measurement scale on ranking a compromise alternative in AHP," *The Journal of the Operational Research Society*, vol. 62, pp. 700-710, 2011.
- [52] S. Siraj, *et al.*, "A heuristic method to rectify intransitive judgments in pairwise comparison matrices," *European Journal of Operational Research*, vol. 216, pp. 420-428, 2012.
- [53] F. Zahedi, "The Analytic Hierarchy Process: A Survey of the Method and Its Applications," *Interfaces*, vol. 16, pp. 96-108, 1986.
- [54] Y. Dong, *et al.*, "A comparative study of the numerical scales and the prioritization methods in AHP," *European Journal of Operational Research*, vol. 186, pp. 229-242, 2008.
- [55] A. Khan, "Capital Rationing, Priority Setting, and Budget Decisions: An Analytical Guide for Public Managers," in *Public Budgeting and Finance, Fourth Edition*, R. T. Golembiewski and J. Rabin, Eds., ed: Taylor & Francis, 1997.
- [56] N. Gerdri, "An Analytical Approach to Building a Technology Development Envelope (TDE) for Roadmapping of Emerging Technologies," *International Journal of Innovation and Technology Management*, vol. 4, pp. 121-135, 2007.
- [57] H. Chen and D. F. Kocaoglu, "A sensitivity analysis algorithm for hierarchical decision models," *European Journal of Operational Research*, vol. 185, pp. 266-288, 2008.
- [58] H. Chen, *et al.*, "A Strategic Technology Planning Framework: A Case of Taiwan's Semiconductor Foundry Industry," *IEEE Transactions on Engineering Management*, vol. 56, pp. 4-15, 2009.
- [59] P. Gerdri and D. F. Kocaoglu, "HDM for developing national emerging technology strategy and policy supporting sustainable economy: A case study of nanotechnology for Thailand's agriculture," in *2008 Portland International Center for Management of Engineering and Technology, Technology Management for a Sustainable Economy, PICMET '08, July 27, 2008 - July 31, 2008*, Cape Town, South africa, 2008, pp. 344-350.

- [60] D. Fenwick, *et al.*, "Value Driven Technology Road Mapping (VTRM) process integrating decision making and marketing tools: Case of Internet security technologies," *Technological Forecasting and Social Change*, vol. 76, pp. 1055-1077, 2009.
- [61] R. Kodali, *et al.*, "Justification of world-class maintenance systems using analytic hierarchy constant sum method," *Journal of Quality in Maintenance Engineering*, vol. 15, pp. 47-77, 2009.
- [62] K. Cowan, *et al.*, "Exploring the impact of technology development and adoption for sustainable hydroelectric power and storage technologies in the Pacific Northwest United States," *Energy*, vol. 35, pp. 4771-4779, 2010.
- [63] T. U. Daim, *et al.*, "Identification of energy policy priorities from existing energy portfolios using hierarchical decision model and goal programming: Case of Germany and France," *International Journal of Energy Sector Management*, vol. 4, pp. 24-43, 2010.
- [64] G. Harell and T. U. Daim, "HDM Modeling as a Tool to Assist Management With Employee Motivation: The Case of Silicon Forest," *Engineering Management Journal*, vol. 22, pp. 23-33, 2010.
- [65] E. Kennedy and T. U. Daim, "A strategy to assist management in workforce engagement and employee retention in the high tech engineering environment," *Evaluation and Program Planning*, vol. 33, pp. 468-476, 2010.
- [66] B. Wang, *et al.*, "A decision model for energy resource selection in China," *Energy Policy*, vol. 38, pp. 7130-7141, 2010.
- [67] D. Fenwick and T. U. Daim, "Choosing a hybrid car using a hierarchical decision model," *International Journal of Sustainable Society*, vol. 3, pp. 243-257, 2011.
- [68] N. Sheikh, *et al.*, "Use of multiple perspectives and decision modeling for PV technology assessment," in *Technology Management in the Energy Smart World (PICMET), 2011 Proceedings of PICMET '11:*, 2011, pp. 1-21.
- [69] T. U. Daim, *et al.*, "Site selection for a data centre – a multi-criteria decision-making model," *International Journal of Sustainable Engineering*, vol. 6, pp. 10-22, 2013/03/01 2012.
- [70] B.-H. Lee, *et al.*, "Relative Priority Analysis of Korean IS Audit Standard Check Items Using the Constant-Sum Method," in *Network and Parallel Computing*. vol. 7513, J. Park, A. Zomaya, S.-S. Yeo, and S. Sahni, Eds., ed: Springer Berlin Heidelberg, 2012, pp. 233-240.
- [71] B. L. Golden and Q. Wang, "An alternate measure of consistency," in *The Analytic Hierarchy Process, Applications and Studies*, B. Golden, E. Wasil, and P. Harker, Eds., ed New York: Springer-Verlag 1990.
- [72] E. F. Lane and W. A. Verdini, "A Consistency Test for AHP Decision Makers," *Decision Sciences*, vol. 20, pp. 575-590, 1989.
- [73] P. T. Harker, "Alternative modes of questioning in the analytic hierarchy process," *Mathematical Modelling*, vol. 9, pp. 353-360, 1987.

- [74] F. J. Dodd, *et al.*, "Inverse inconsistency in analytic hierarchies," *European Journal of Operational Research*, vol. 80, pp. 86-93, 1995.
- [75] V. M. R. Tummala and Y.-w. Wan, "On the mean random inconsistency index of analytic hierarchy process (AHP)," *Computers and Industrial Engineering*, vol. 27, pp. 401-404, 1994.
- [76] W. W. Koczkodaj, "A new definition of consistency of pairwise comparisons," *Mathematical and Computer Modelling*, vol. 18, pp. 79-84, 1993.
- [77] J. Barzilai, *et al.*, "Consistent weights for judgements matrices of the relative importance of alternatives," *Operations Research Letters*, vol. 6, pp. 131-134, 1987.
- [78] G. Crawford and C. Williams, "A note on the analysis of subjective judgment matrices," *Journal of Mathematical Psychology*, vol. 29, pp. 387-405, 1985.
- [79] L. G. Vargas, "Reciprocal matrices with random coefficients," *Mathematical Modelling*, vol. 3, pp. 69-81, 1982.
- [80] G. A. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information," *Psychological Review*, vol. 63, pp. 81-97, 1956.
- [81] S. Karapetrovic and E. S. Rosenbloom, "Quality control approach to consistency paradoxes in AHP," *European Journal of Operational Research*, vol. 119, pp. 704-718, 1999.
- [82] A. A. Salo and R. P. Hämäläinen, "On the measurement of preferences in the analytic hierarchy process," *Journal of Multi-Criteria Decision Analysis*, vol. 6, pp. 309-319, 1997.
- [83] C. A. Bana e Costa and J.-C. Vansnick, "A critical analysis of the eigenvalue method used to derive priorities in AHP," *European Journal of Operational Research*, vol. 187, pp. 1422-1428, 2008.
- [84] J. Barzilai and B. Golany, "Deriving weights from pairwise comparison matrices: The additive case," *Operations Research Letters*, vol. 9, pp. 407-410, 1990.
- [85] J. W. Ra, "Hierarchical decision process," in *Technology Management : the New International Language*, 1991, pp. 595-599.
- [86] G. Kretchik and J. W. Ra, "Analysis of the sum of the inverse column sums (SICS): an alternative consistency measure for pairwise comparisons," in *Management of Engineering and Technology, 1999. Technology and Innovation Management. PICMET '99. Portland International Conference on Management of Engineering and Technology*, 1999, p. 229 vol.1.
- [87] R. E. Jensen and T. E. Hicks, "Ordinal data AHP analysis: A proposed coefficient of consistency and a nonparametric test," *Mathematical and Computer Modelling*, vol. 17, pp. 135-150, 1993.
- [88] E. Takeda, "A note on consistent adjustments of pairwise comparison judgments," *Mathematical and Computer Modelling*, vol. 17, pp. 29-35, 1993.
- [89] W. C. Wedley, "Consistency prediction for incomplete AHP matrices," *Mathematical and Computer Modelling*, vol. 17, pp. 151-161, 1993.

- [90] E. Takeda and P.-L. Yu, "Assessing priority weights from subsets of pairwise comparisons in multiple criteria optimization problems," *European Journal of Operational Research*, vol. 86, pp. 315-331, 1995.
- [91] H. Monsuur, "An intrinsic consistency threshold for reciprocal matrices," *European Journal of Operational Research*, vol. 96, pp. 387-391, 1997.
- [92] J. Barzilai, "Consistency measures for pairwise comparison matrices," *Journal of Multi-Criteria Decision Analysis*, vol. 7, pp. 123-132, 1998.
- [93] S. Shiraishi, *et al.*, "Properties of a Positive Reciprocal Matrix and their Application to AHP," *Operations Research Society of Japan*, vol. 41, pp. 404-414, 1998.
- [94] J. I. Peláez and M. T. Lamata, "A new measure of consistency for positive reciprocal matrices," *Computers & Mathematics with Applications*, vol. 46, pp. 1839-1845, 2003.
- [95] S. I. Gass and T. Rapcsák, "Singular value decomposition in AHP," *European Journal of Operational Research*, vol. 154, pp. 573-584, 2004.
- [96] M. Fedrizzi and S. Giove, "Incomplete pairwise comparison and consistency optimization," *European Journal of Operational Research*, vol. 183, pp. 303-313, 2007.
- [97] W. E. Stein and P. J. Mizzi, "The harmonic consistency index for the analytic hierarchy process," *European Journal of Operational Research*, vol. 177, pp. 488-497, 2007.
- [98] M. Fedrizzi and M. Brunelli, "Fair Consistency Evaluation for Reciprocal Relations and in Group Decision Making," *New Mathematics & Natural Computation*, vol. 5, pp. 407-420, 2009.
- [99] L. Čaklović, "Measure of Inconsistency for the Potential Method," in *Modeling Decisions for Artificial Intelligence*. vol. 7647, V. Torra, Y. Narukawa, B. López, and M. Villaret, Eds., ed: Springer Berlin / Heidelberg, 2012, pp. 102-114.
- [100] B. Matteo, *et al.*, "A note on the proportionality between some consistency indices in the AHP," *CORD Conference Proceedings*, 2012.
- [101] S. Siraj, *et al.*, "Enumerating all spanning trees for pairwise comparisons," *Computers & Operations Research*, vol. 39, pp. 191-199, 2012.
- [102] F. J. Dodd, *et al.*, "A statistical approach to consistency in AHP," *Mathematical and Computer Modelling*, vol. 18, pp. 19-22, 1993.
- [103] G. Kou, *et al.*, *Data Processing for the AHP/ANP*: Springer, 2013.
- [104] J. D. Gibbons, *Nonparametric Methods for Quantitative Analysis*: American Sciences Press, 1985.
- [105] A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis*, Third ed.: McGraw-Hill Higher Education, 2000.
- [106] K. Beven, *Environmental Modelling: An Uncertain Future?*: Taylor & Francis, 2010.
- [107] R. Korn, *et al.*, *Monte Carlo Methods and Models in Finance and Insurance*: CRC Press, 2010.

- [108] C. Cioffi-Revilla, "On the Quality of a Social Simulation Model: A Lifecycle Framework," in *Advances in Social Simulation*. vol. 229, B. Kamiński and G. Koloch, Eds., ed: Springer Berlin Heidelberg, 2014, pp. 13-23.
- [109] W. Vanderbauwhede and K. Benkrid, *High-Performance Computing Using FPGAs*: Springer, 2013.
- [110] M. Manssen, *et al.*, "Random number generators for massively parallel simulations on GPU," *The European Physical Journal Special Topics*, vol. 210, pp. 53-71, 2012/08/01 2012.
- [111] A. S. K. Pathan, *et al.*, *Simulation Technologies in Networking and Communications: Selecting the Best Tool for the Test*: Taylor & Francis, 2014.
- [112] M. S. Obaidat and N. A. Boudriga, *Fundamentals of Performance Evaluation of Computer and Telecommunications Systems*: Wiley, 2010.
- [113] T. D. Little, *The Oxford Handbook of Quantitative Methods, Volume 1: Foundations*: Oxford University Press, 2013.
- [114] S. Goldshtein, *et al.*, *Pro .NET Performance: Optimize Your C# Applications*: Apress, 2012.
- [115] L. Gherardi, *et al.*, "A Java vs. C++ Performance Evaluation: A 3D Modeling Benchmark," in *Simulation, Modeling, and Programming for Autonomous Robots*. vol. 7628, I. Noda, N. Ando, D. Brugali, and J. Kuffner, Eds., ed: Springer Berlin Heidelberg, 2012, pp. 161-172.
- [116] S. B. Aruoba and J. Fernández-Villaverde, "A Comparison of Programming Languages in Economics," *National Bureau of Economic Research Working Paper Series*, vol. No. 20263, 2014.
- [117] K. Lange, *Numerical Analysis for Statisticians*: Springer New York, 2010.
- [118] S. G. Whittington, *Numerical Methods for Polymeric Systems*: Springer New York, 2012.
- [119] A. Gosavi, *Simulation-Based Optimization: Parametric Optimization Techniques and Reinforcement Learning*: Springer US, 2014.
- [120] N. Karamanis, *et al.*, "A corpus-based methodology for evaluating metrics of coherence for text structuring," in *Natural Language Generation*, ed: Springer, 2004, pp. 90-99.
- [121] N. Cressie, *Statistics for Spatial Data*: Wiley, 2015.
- [122] W. Trochim, *et al.*, *Research Methods: The Essential Knowledge Base*: Cengage Learning, 2015.
- [123] R. B. Johnson and L. Christensen, *Educational Research: Quantitative, Qualitative, and Mixed Approaches*: SAGE Publications, 2013.
- [124] C. E. Särndal, *et al.*, *Model Assisted Survey Sampling*: Springer New York, 2003.
- [125] D. H. Lehmer, "Teaching combinatorial tricks to a computer," *Proc. Sympos. Appl. Math. Combinatorial Analysis American Mathematical Society, Providence, R.I.*, vol. 10, pp. 179-193, 1960.
- [126] J. L. Gross, *Combinatorial Methods with Computer Applications*: Taylor & Francis, 2007.

- [127] S. Olariu and A. Y. Zomaya, *Handbook of Bioinspired Algorithms and Applications*: CRC Press, 2005.
- [128] M. Roberto and R. Fanja, "A permutations representation that knows what "Eulerian" means," *Discrete Mathematics & Theoretical Computer Science*; Vol 4, No 2 (2001), 2006.
- [129] P. Tarau, "A Groupoid of Isomorphic Data Transformations," in *Intelligent Computer Mathematics*. vol. 5625, J. Carette, L. Dixon, C. Coen, and S. Watt, Eds., ed: Springer Berlin Heidelberg, 2009, pp. 170-185.
- [130] S. Pesko, "Differential Evolution for Small TSPs with Constraints," in *Proceedings of the fourth International Scientific Conference: Challenges in Transport and Communications, Part III*, Pardubice, Czech Republic, 2006, pp. 989–994.
- [131] D. E. Knuth, *The Art of Computer Programming, Volume 3: Sorting and Searching. Second Edition*: Addison-Wesley Longman Publishing Co., Inc., 1997.
- [132] M. Hall, *Proc. Sympos. Appl. Math. Combinatorial Analysis American Mathematical Society, Providence, R.I.*, vol. 6, p. 203 and following, 1956.
- [133] M.-E. Marmion and O. Regnier-Coudert, "Fitness Landscape of the Factoradic Representation on the Permutation Flowshop Scheduling Problem," in *Learning and Intelligent Optimization*. vol. 8994, C. Dhaenens, L. Jourdan, and M.-E. Marmion, Eds., ed: Springer International Publishing, 2015, pp. 151-164.
- [134] H. Samarghandi, *et al.*, "A particle swarm optimization for the single row facility layout problem," *Computers & Industrial Engineering*, vol. 58, pp. 529-534, 2010.
- [135] L. G. Vargas, "The consistency index in reciprocal matrices: Comparison of deterministic and statistical approaches," *European Journal of Operational Research*, vol. 191, pp. 454-463, 2008.
- [136] H. Guerrero, *Excel Data Analysis: Modeling and Simulation*: Springer Berlin Heidelberg, 2010.
- [137] T. L. Saaty and M. S. Ozdemir, "Why the magic number seven plus or minus two," *Mathematical and Computer Modelling*, vol. 38, pp. 233-244, 2003.
- [138] J. W. Ra, "Analysis of Expert Judgments in Hierarchical Decision Process," *PhD Dissertation, University of Pittsburgh*, pp. 118 -128, 1988.



## APPENDIX A: PDF for $n = 3 - 12$

RSV which is HDM's inconsistency measure is assumed to be a continuous random variable. The probability density function (PDF) of RSV defines the magnitude of probability over the entire range of possible RSV values (the RSV domain). There are a few points to keep in mind when using pdf:

- Probabilities are measured over intervals and the integral over an interval is the probability magnitude. Since the variable (RSV) is continuous, so is its pdf. This implies that the probability of an exact value is always zero. If  $X$  denotes particular RSV value for which the probability is to be determined,  $a$  and  $b$  denote the lower and upper limits of  $X$ , and  $f(x)$  denotes RSV's pdf, then the probability that  $X$  falls in  $(a, b]$  is the area under the graph of  $f$  between  $a$  and  $b$  :

$$P\{a < X \leq b\} = \int_a^b f(x)dx$$

for every interval  $(a, b]$ .

- The amplitude of pdf could be higher than one. However, when pdf is integrated over the entire domain of the random variable (RSV in this case), the result should always be one. This unity represents the total probability of all possible values:

$$\int f(x)dx = 1$$

where the integral is implicitly taken over the entire domain of  $X$ .

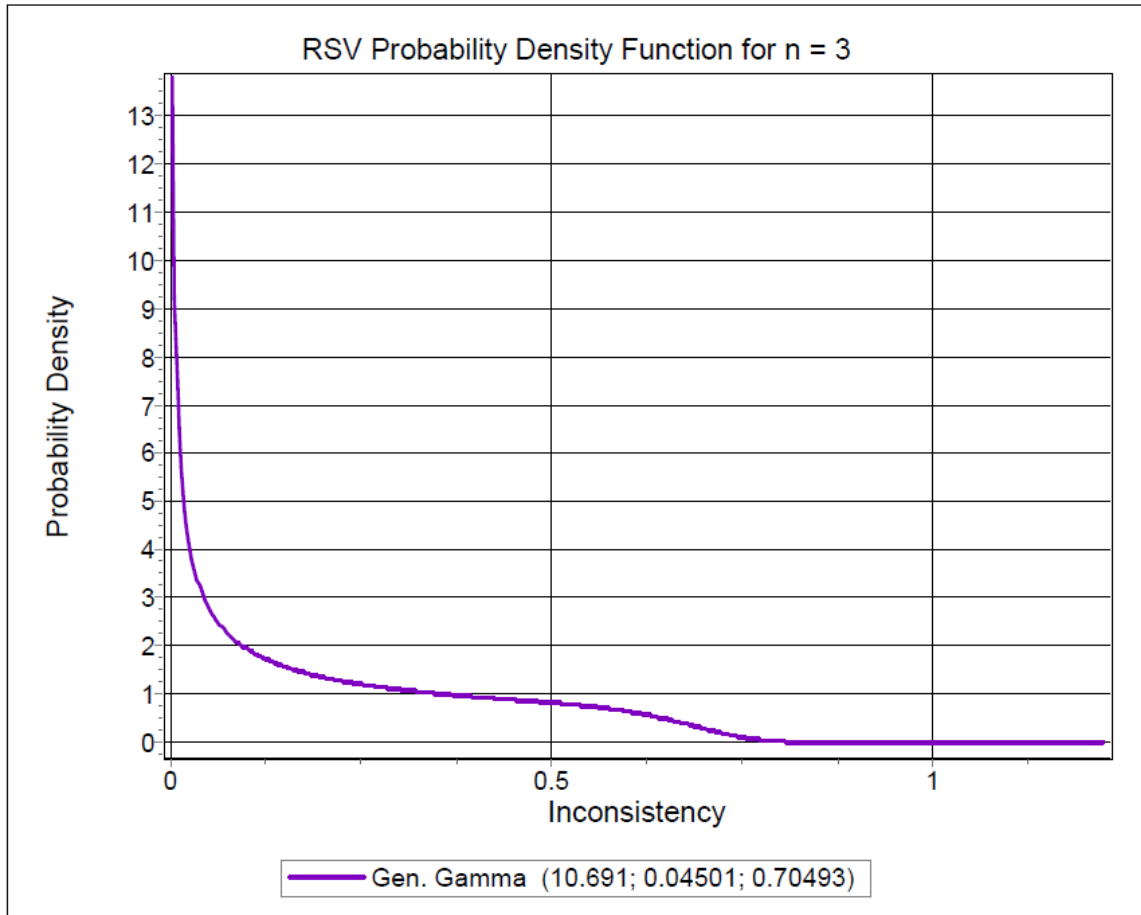


Figure 7: Probability Density Function (PDF) for 3 Variables

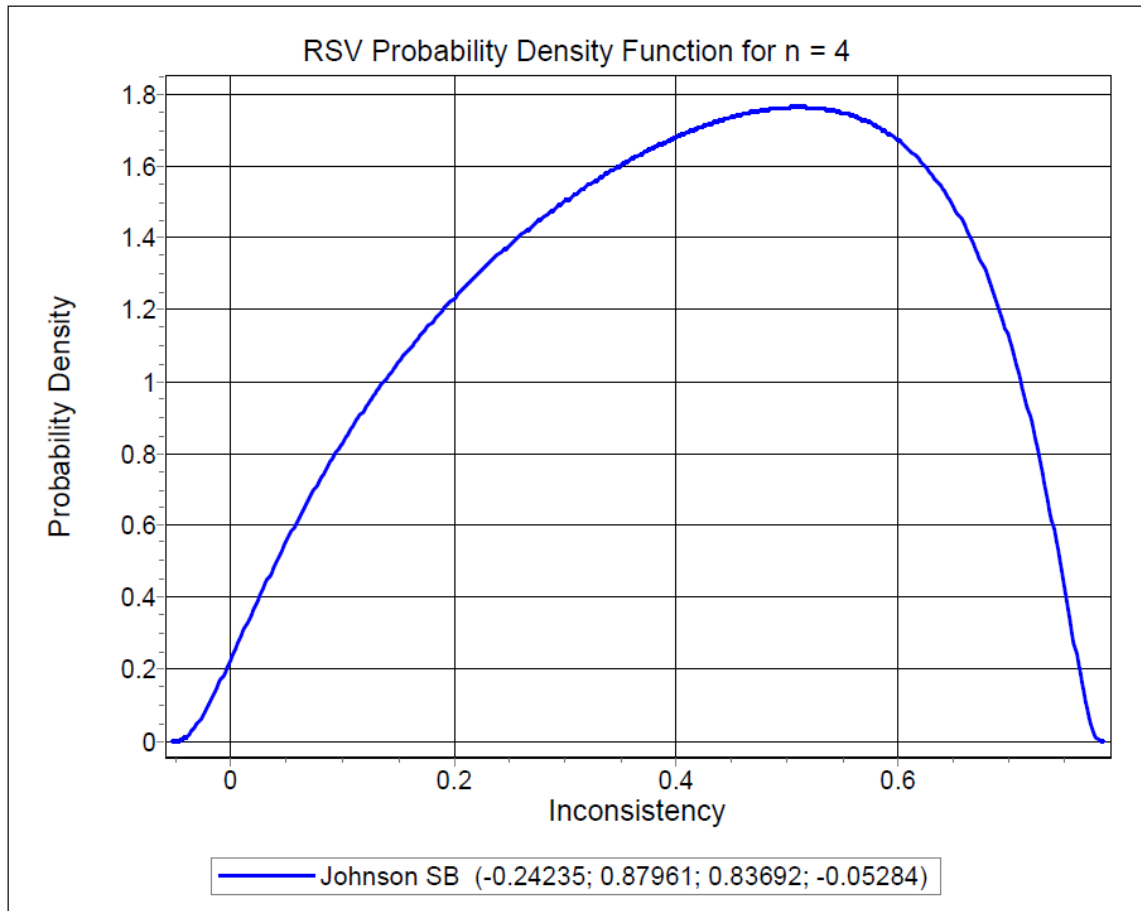


Figure 8: Probability Density Function (PDF) for 4 Variables

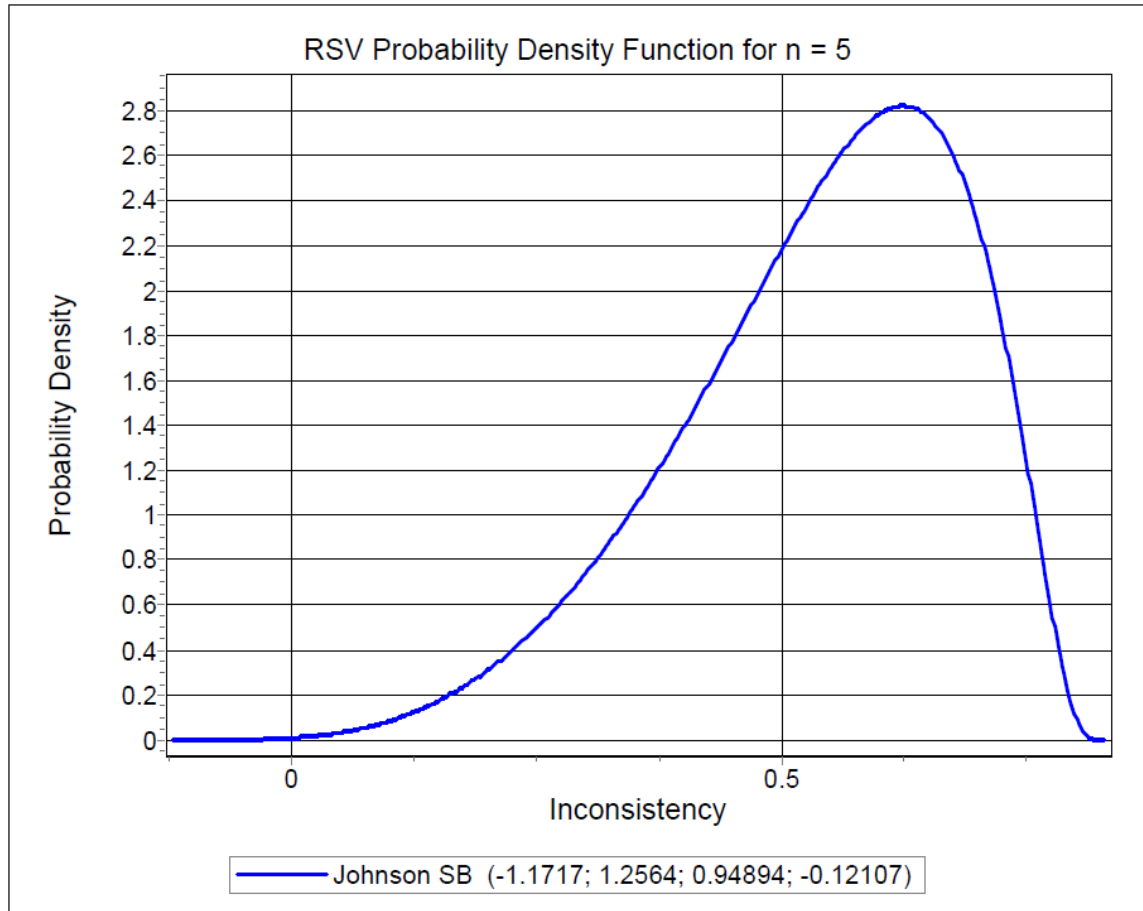


Figure 9: Probability Density Function (PDF) for 5 Variables

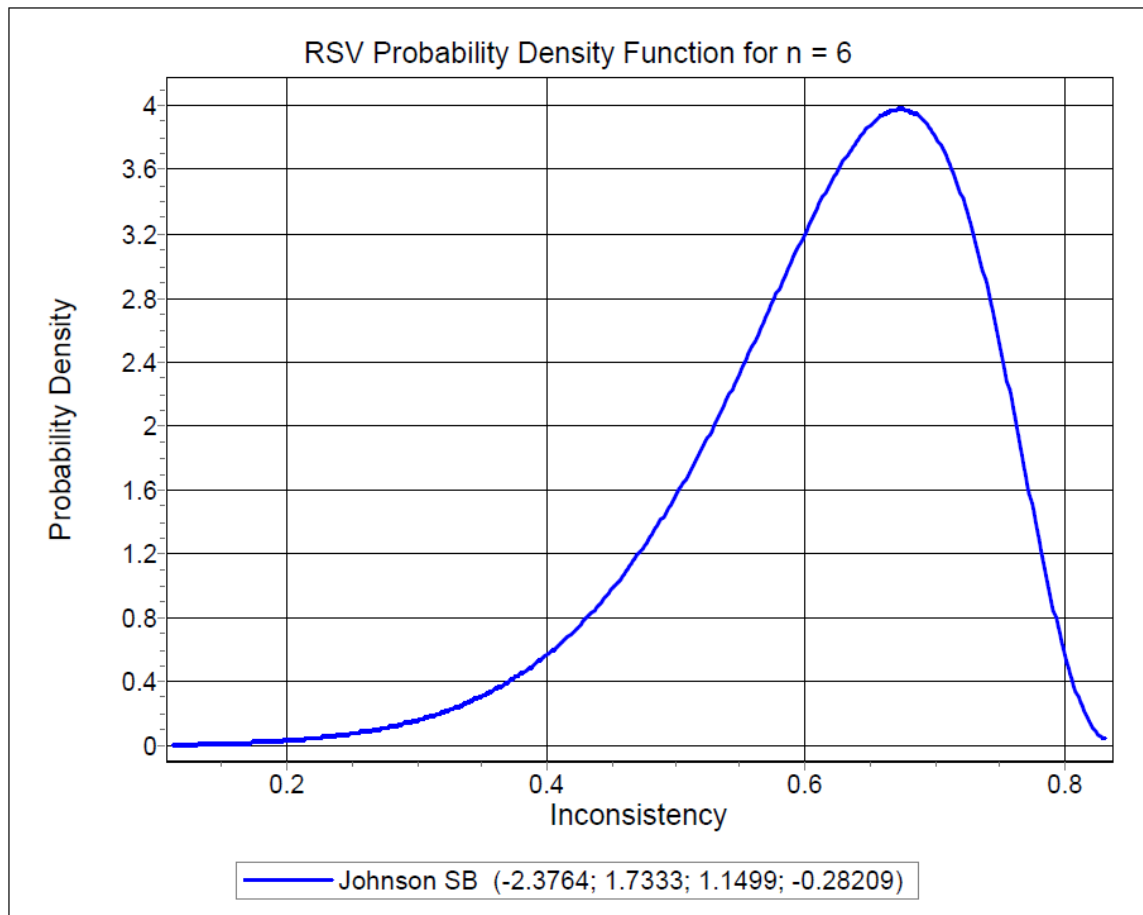


Figure 10: Probability Density Function (PDF) for 6 Variables

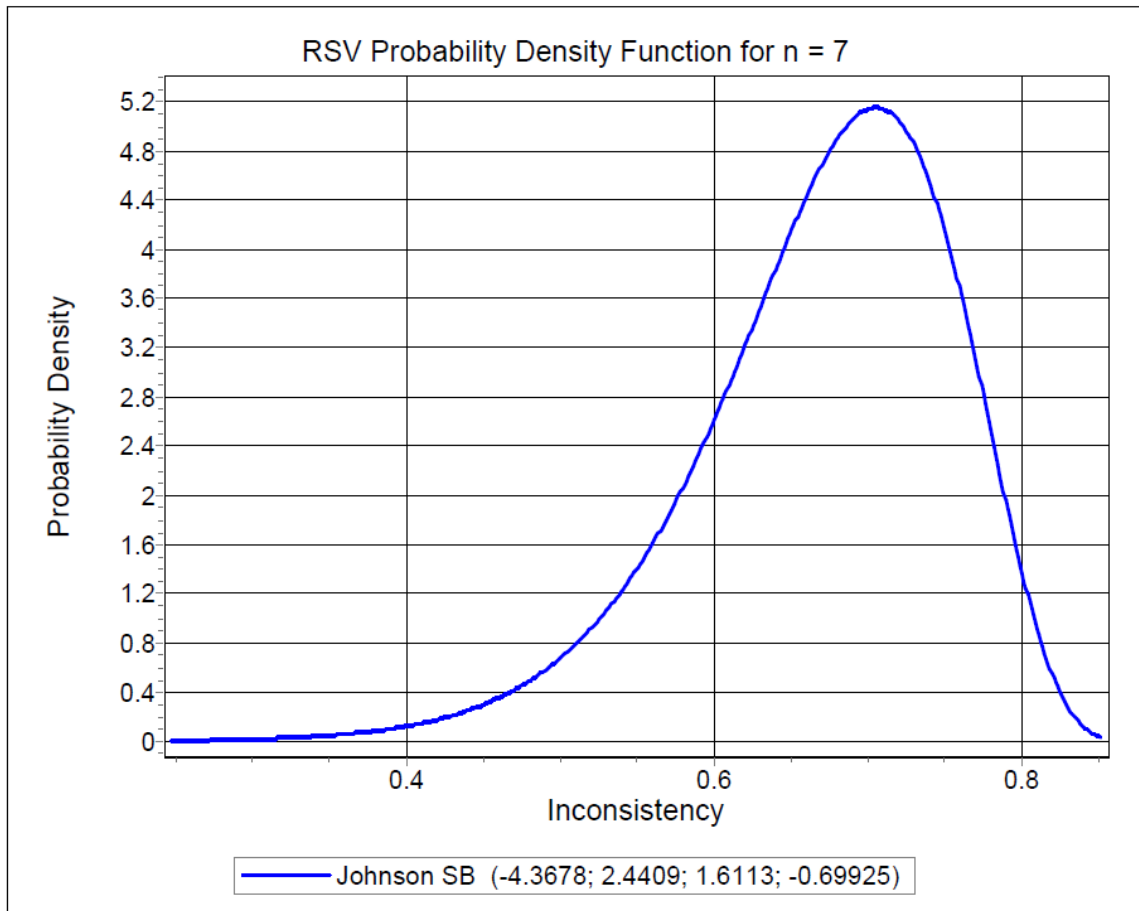


Figure 11: Probability Density Function (PDF) for 7 Variables

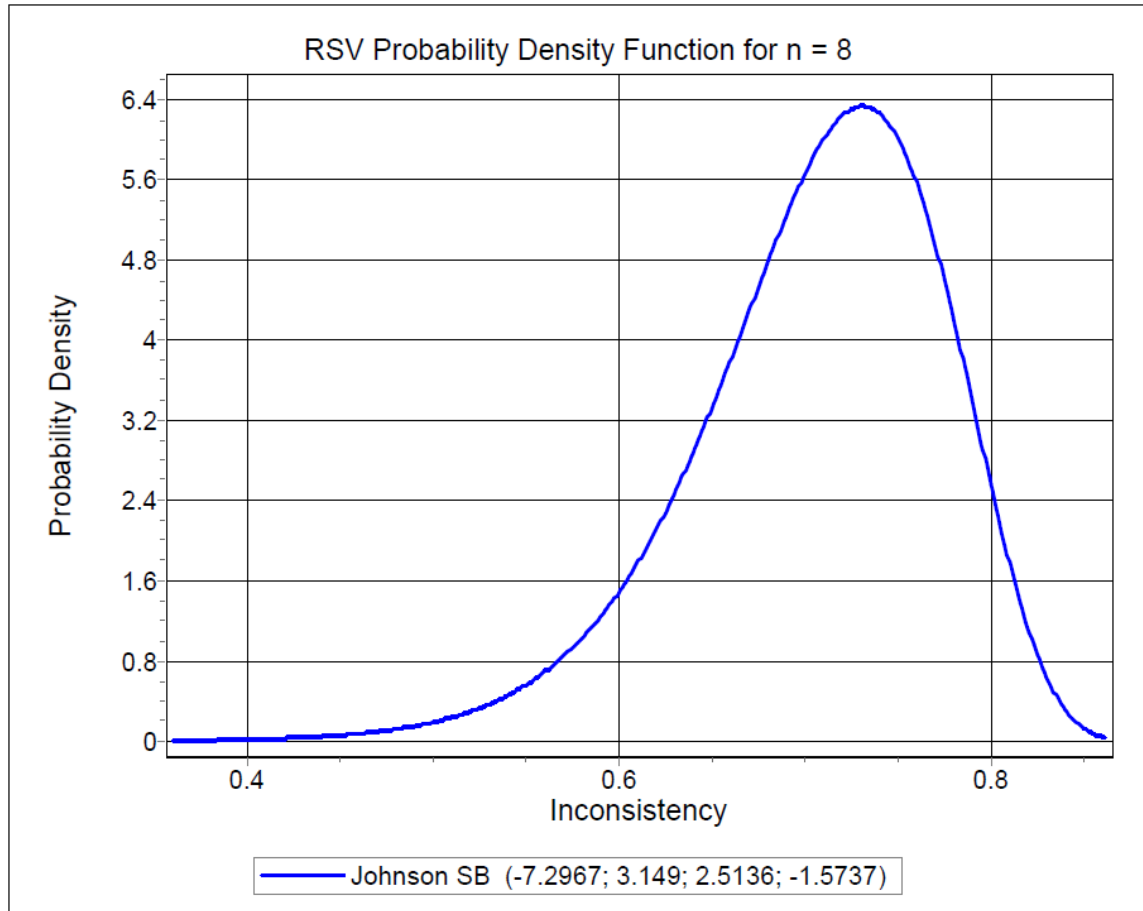


Figure 12: Probability Density Function (PDF) for 8 Variables

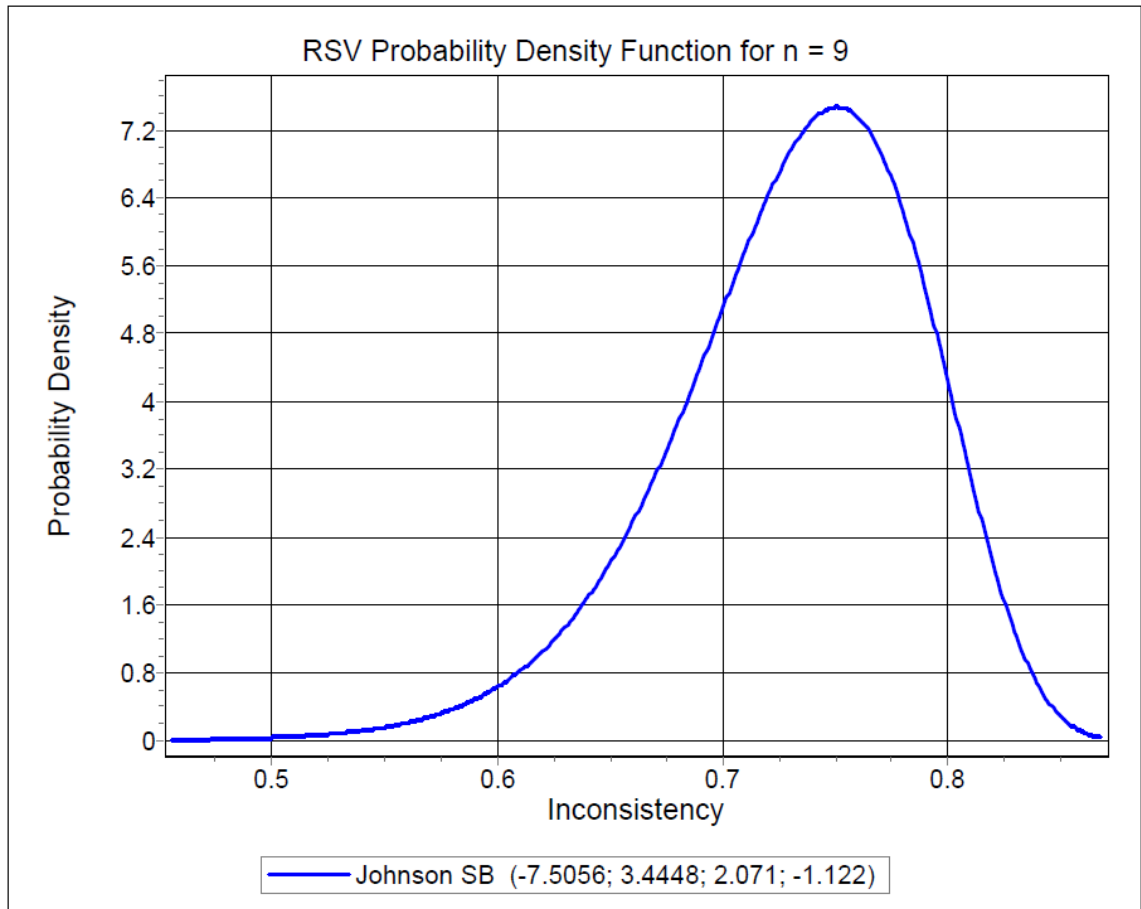


Figure 13: Probability Density Function (PDF) for 9 Variables



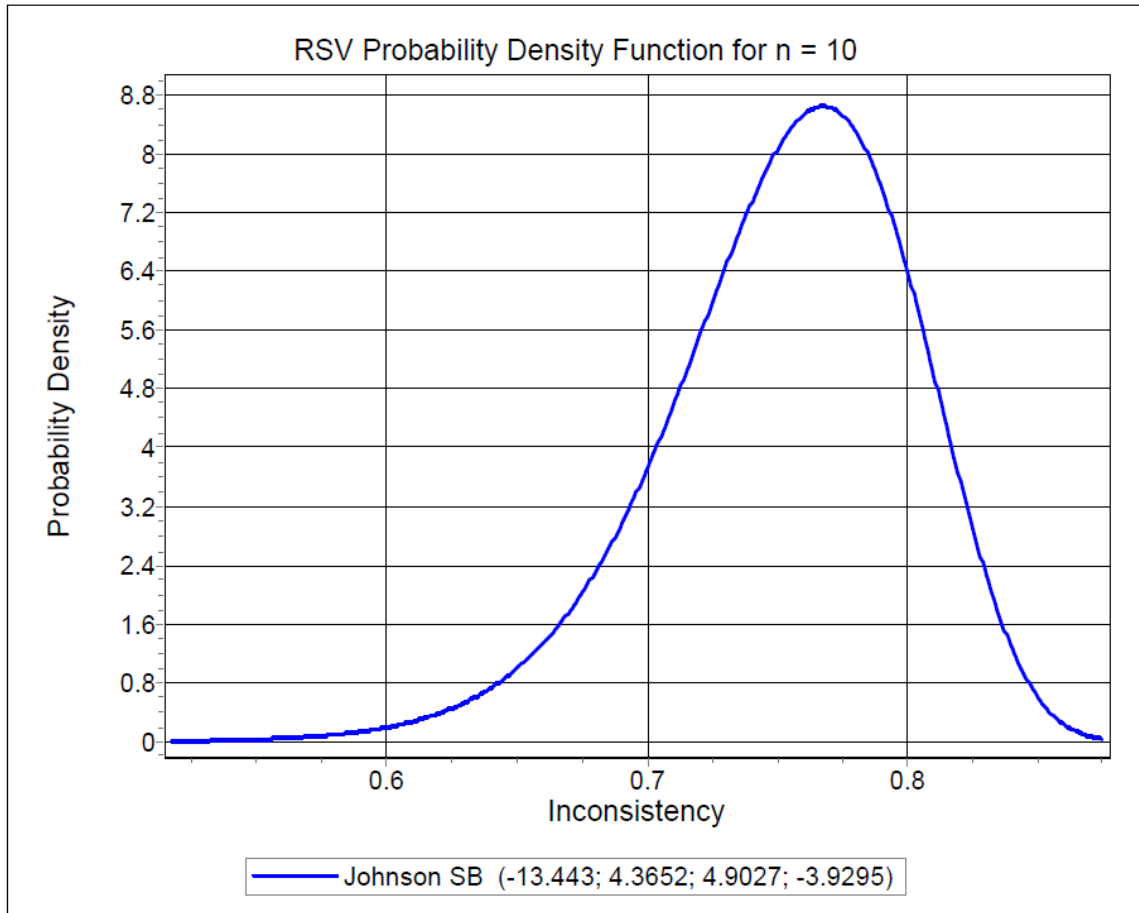


Figure 14: Probability Density Function (PDF) for 10 Variables

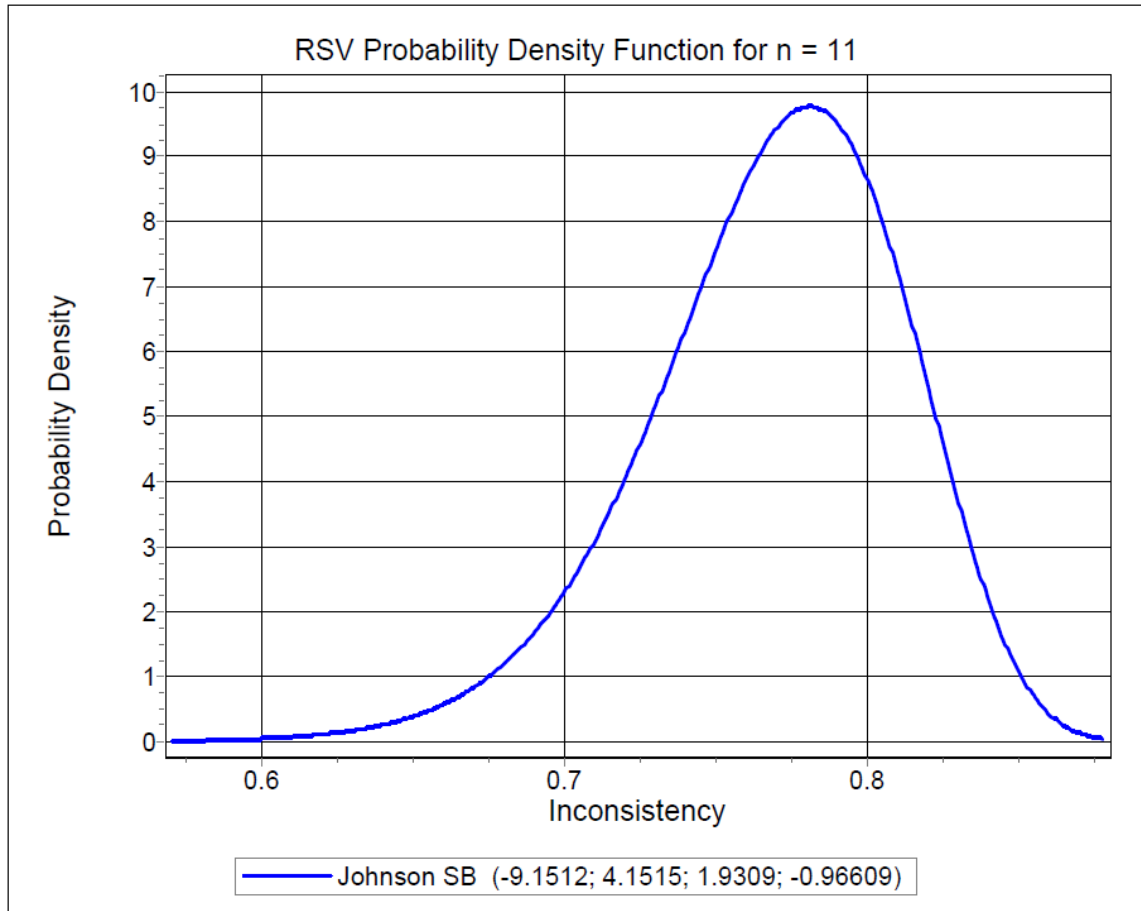


Figure 15: Probability Density Function (PDF) for 11 Variables

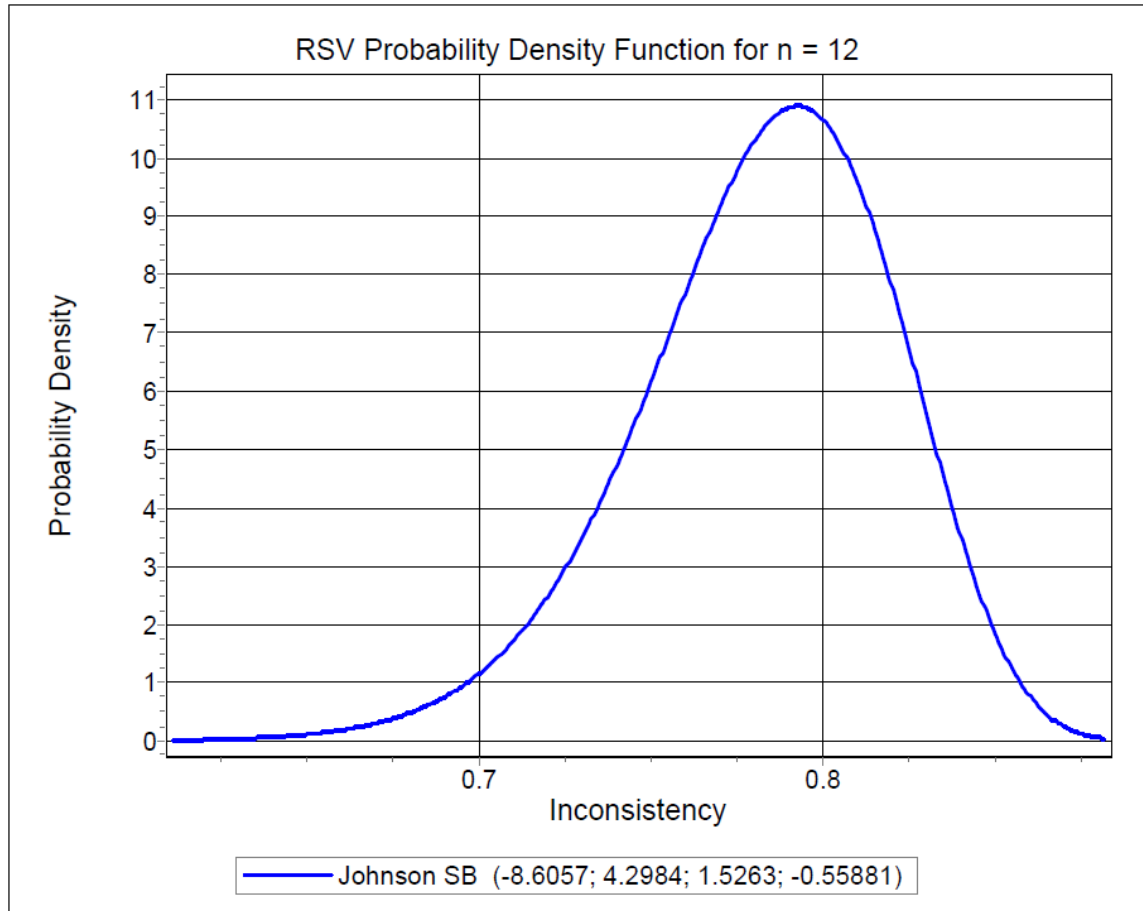


Figure 16: Probability Density Function (PDF) for 12 Variables

## APPENDIX B: CDF for $n = 3 - 12$

A percentile is a statistic that defines a value below which a given percentage of the domain (possible values of a random variable) occurs. The function that maps the percentiles to the domain is the cumulative distribution function (CDF). If  $X$  denotes particular RSV value for which the percentile is to be determined,  $f(x)$  denotes RSV's pdf, and  $F(x)$  denotes RSV's CDF, then:

$$F(x) = \int_{-\infty}^x f(t)dt$$

There are a few points to keep in mind when using pdf:

- In the case of HDM's inconsistency, percentiles are considered the inconsistency thresholds (limits).
- Since inconsistency is a measure of error, the lower the percentile, the more consistent the pairwise comparisons and the more sound the decision. This is the same as the  $\alpha$  level

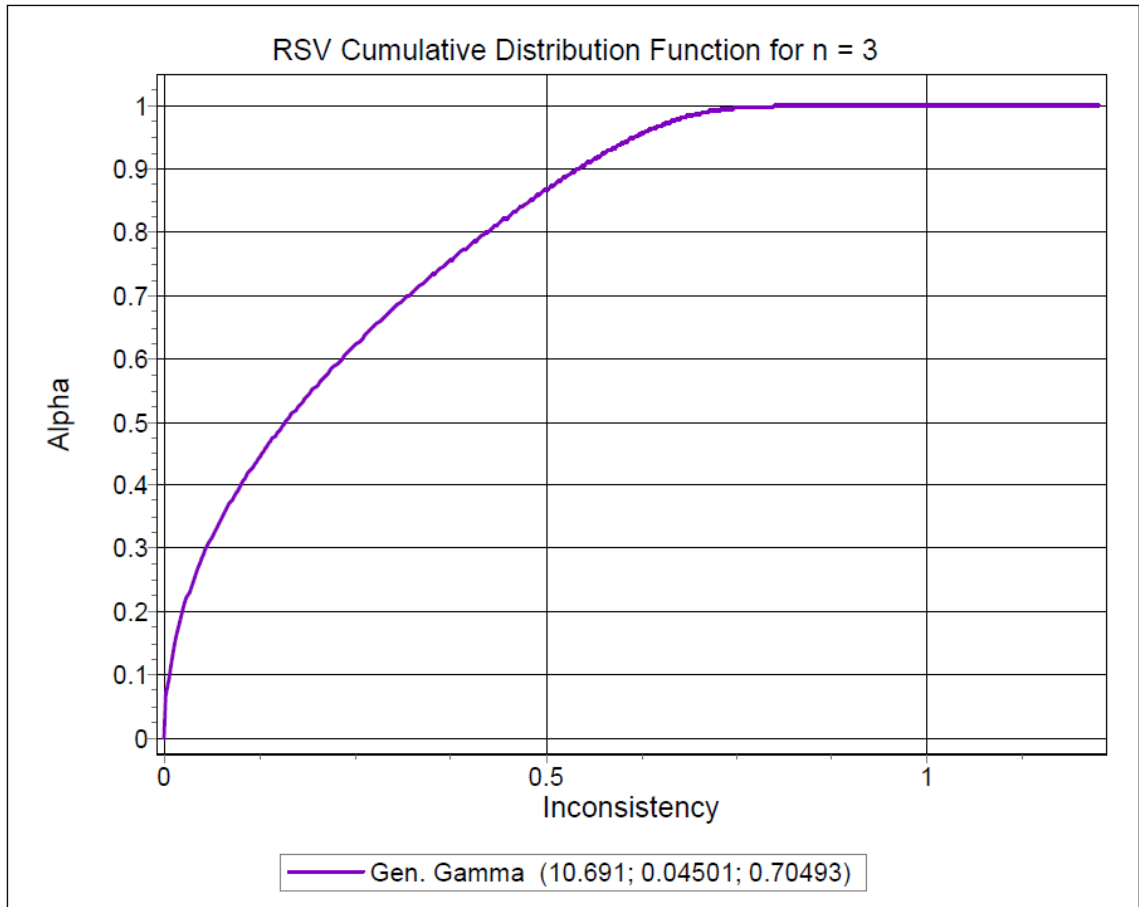


Figure 17: Cumulative Distribution Function (CDF) for 3 Variables

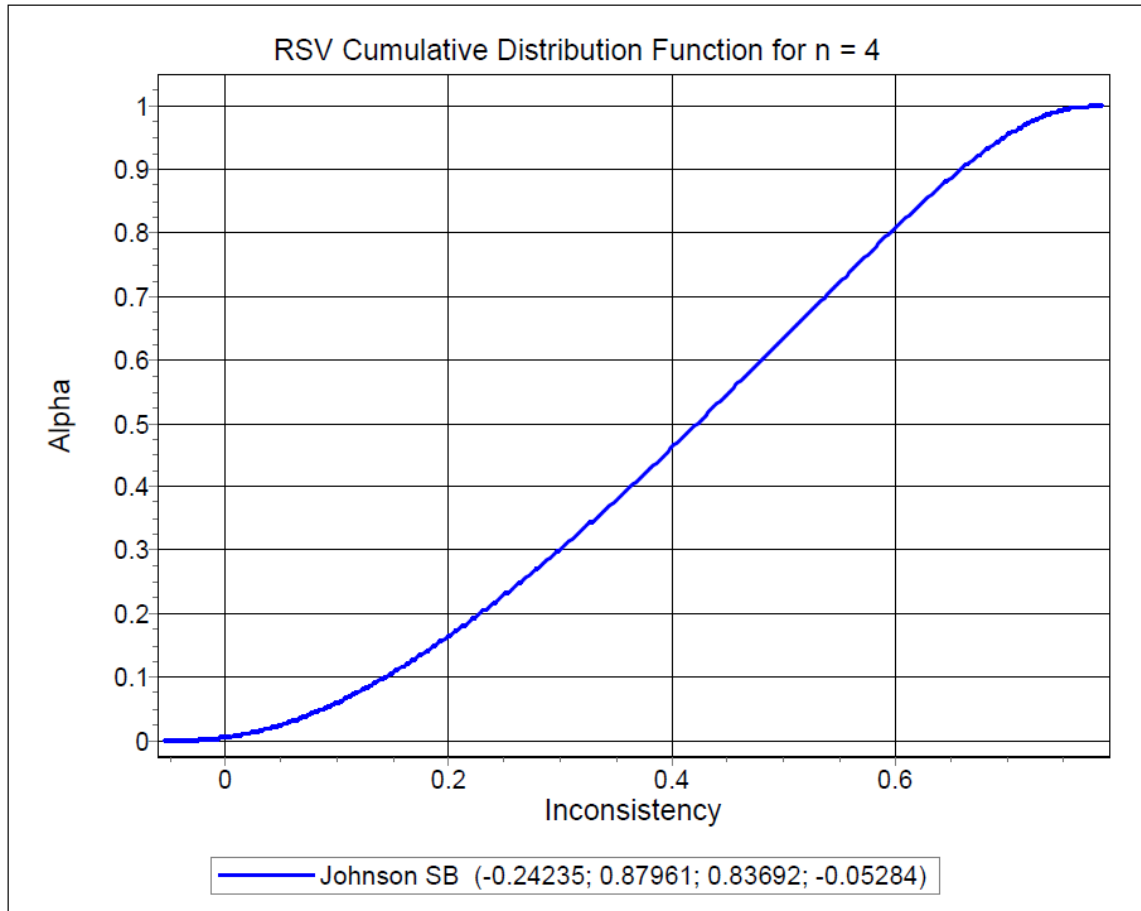


Figure 18: Cumulative Distribution Function (CDF) for 4 Variables

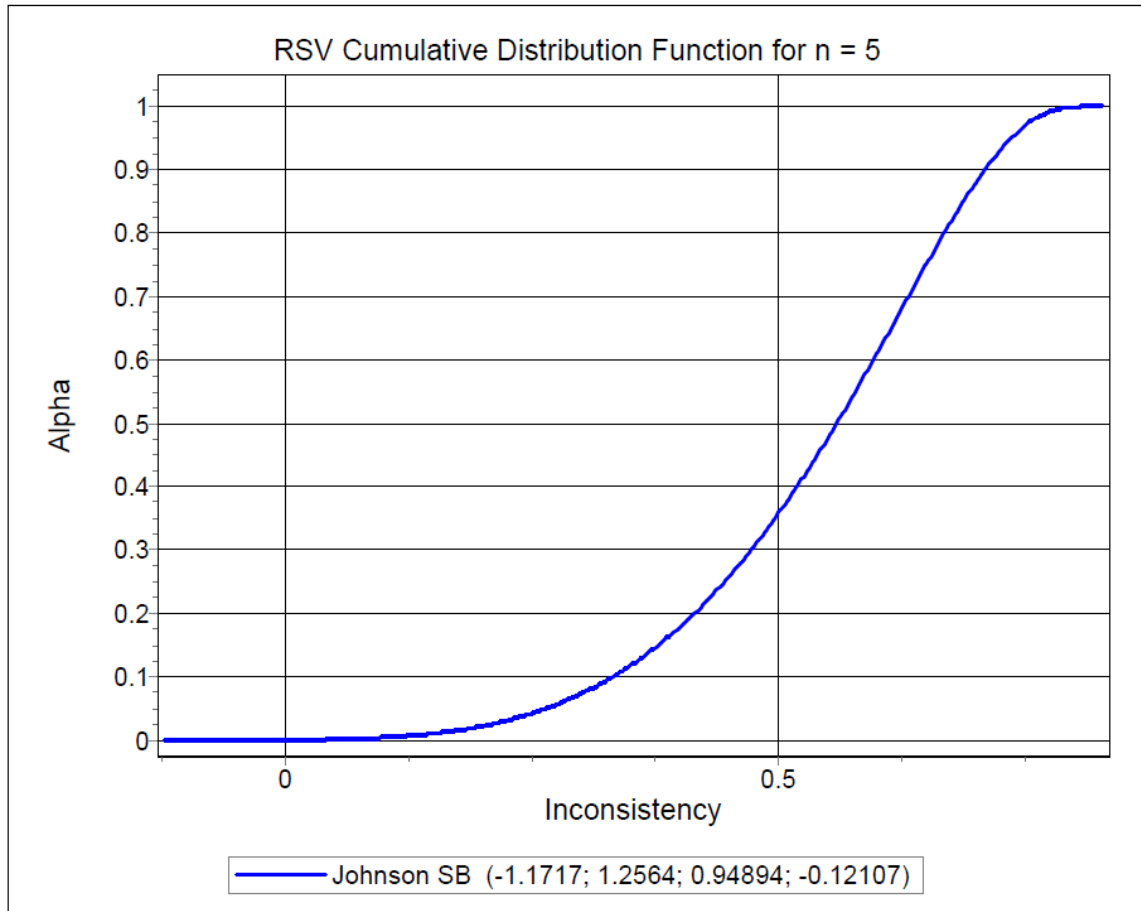


Figure 19: Cumulative Distribution Function (CDF) for 5 Variables

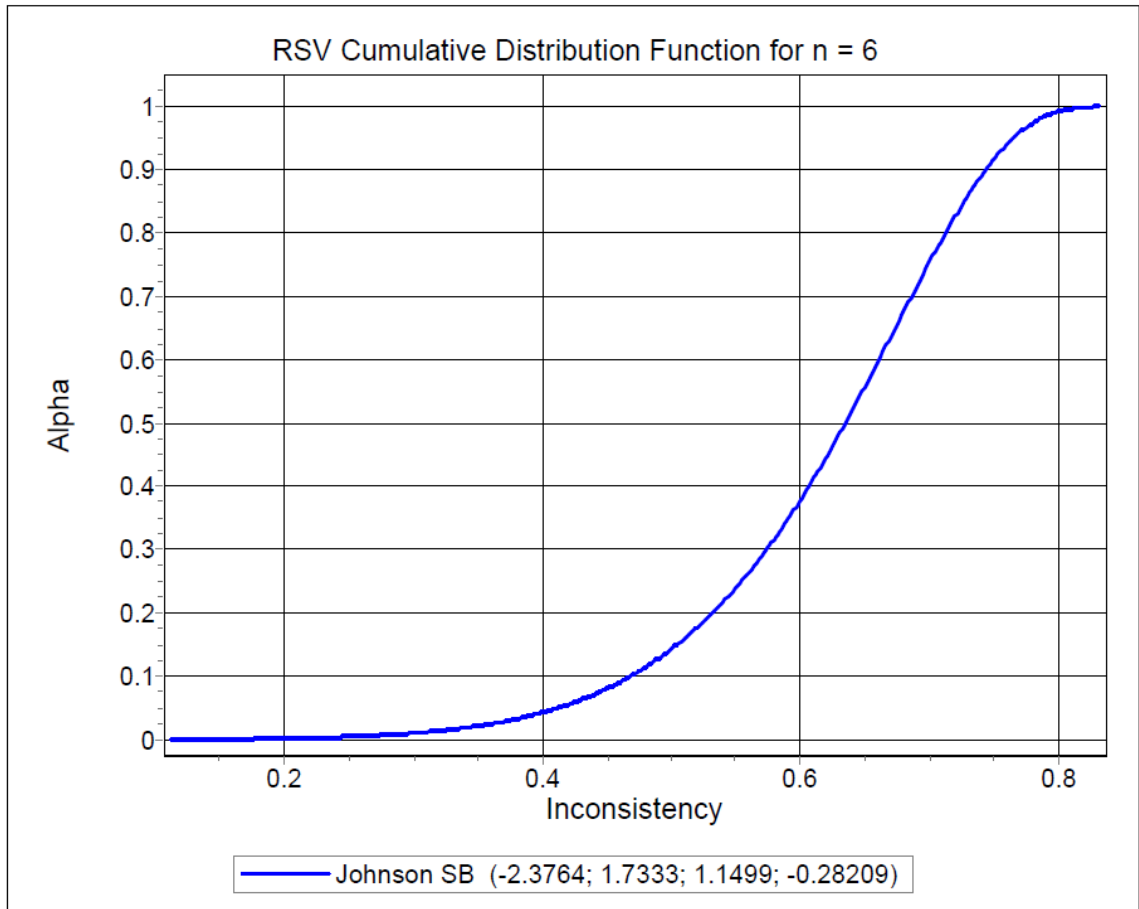


Figure 20: Cumulative Distribution Function (CDF) for 6 Variables



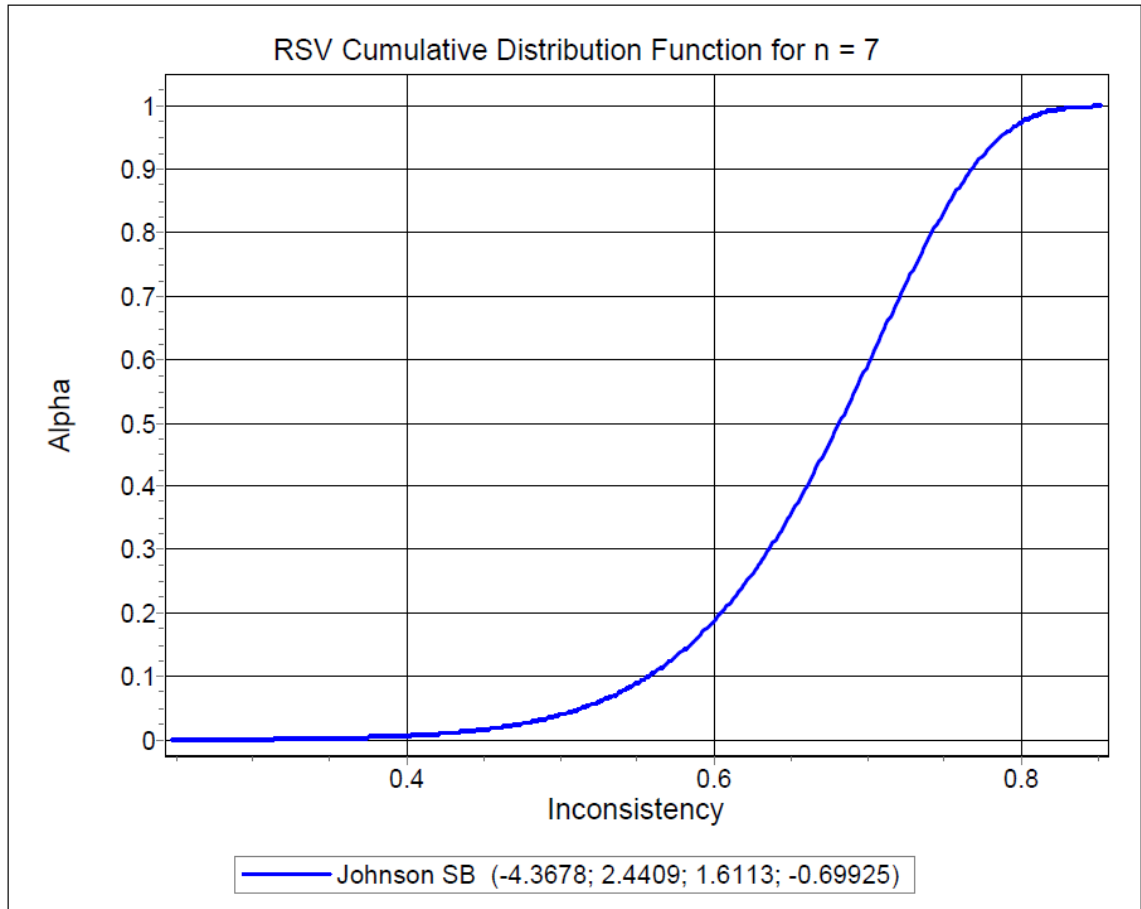


Figure 21: Cumulative Distribution Function (CDF) for 7 Variables

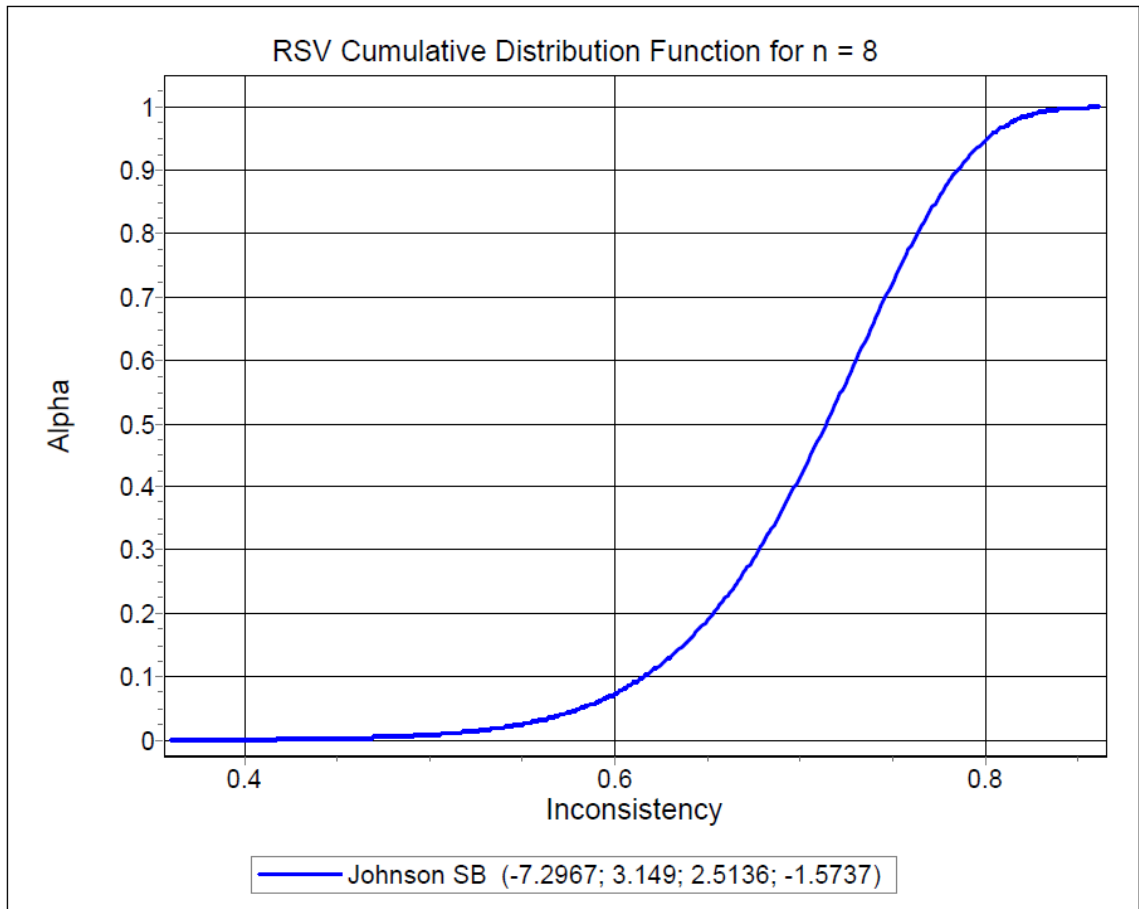


Figure 22: Cumulative Distribution Function (CDF) for 8 Variables

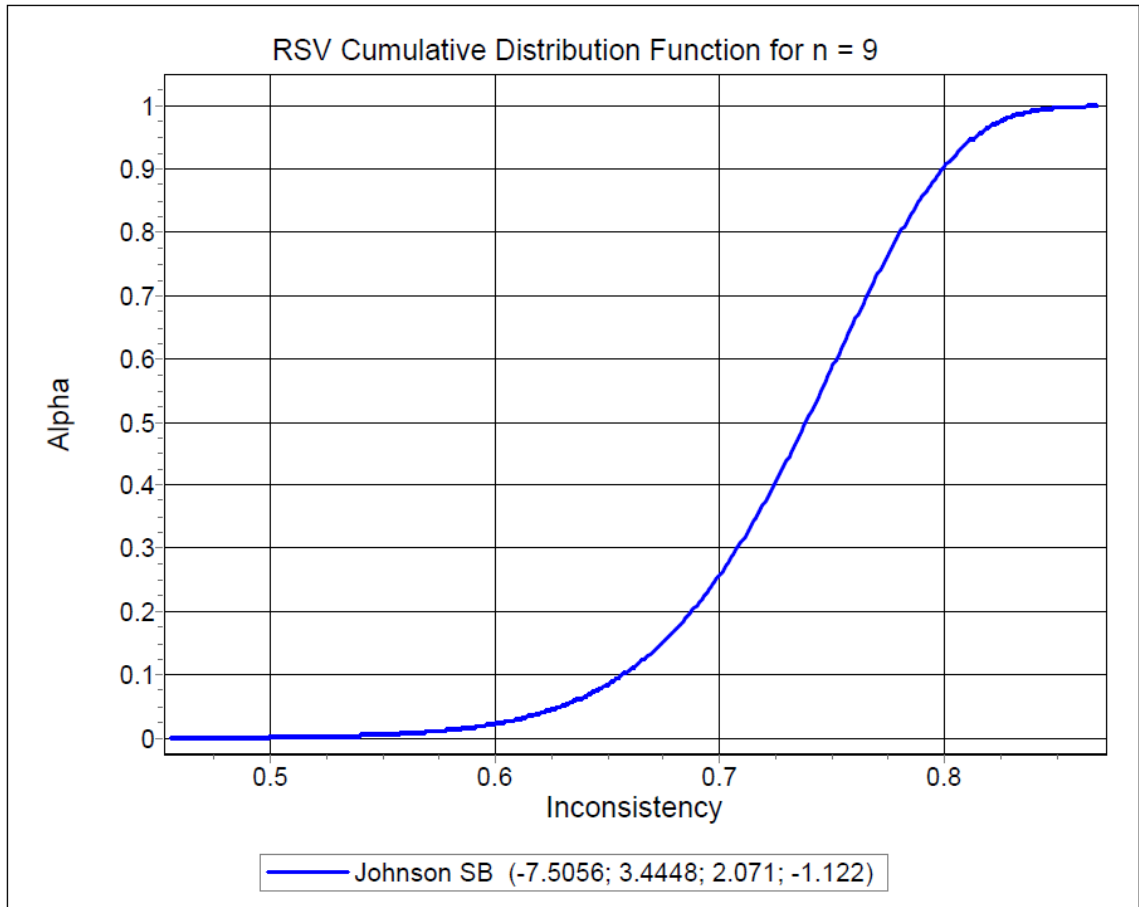


Figure 23: Cumulative Distribution Function (CDF) for 9 Variables

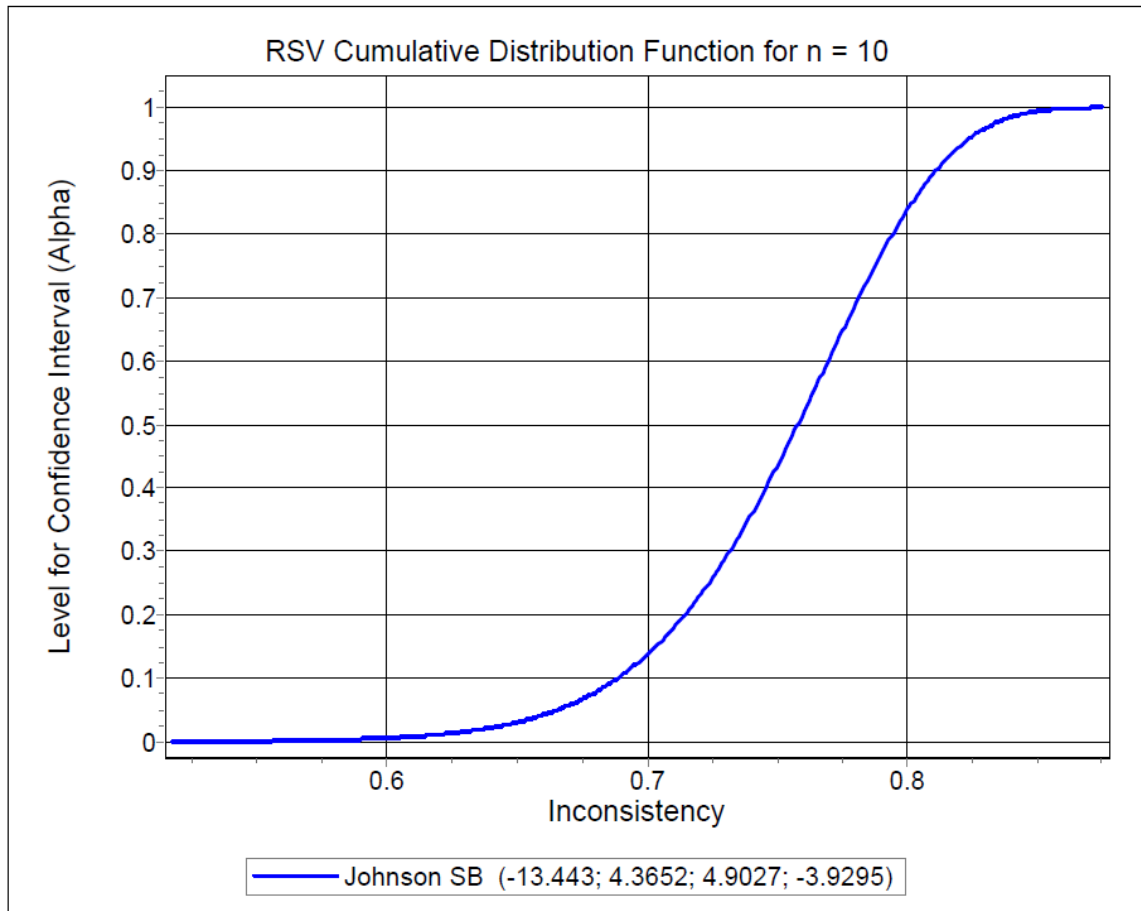


Figure 24: Cumulative Distribution Function (CDF) for 10 Variables

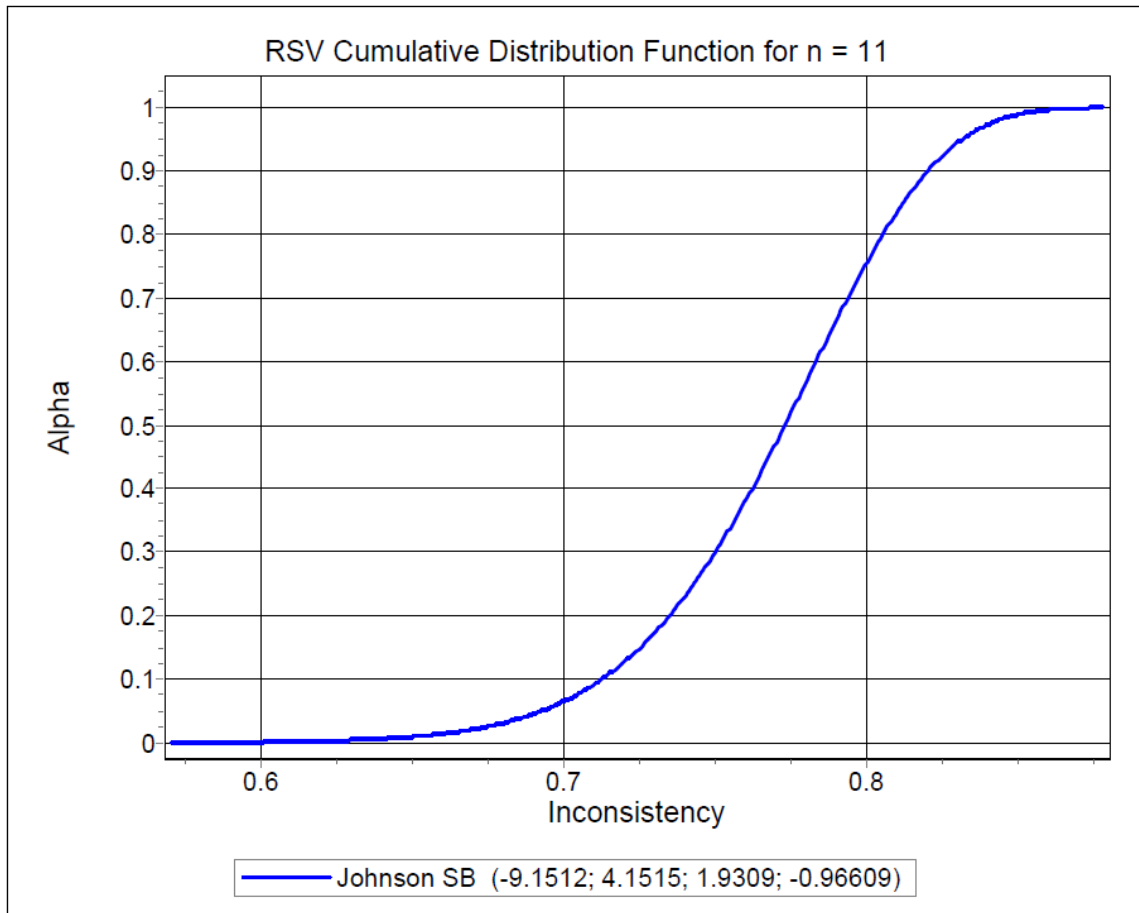


Figure 25: Cumulative Distribution Function (CDF) for 11 Variables

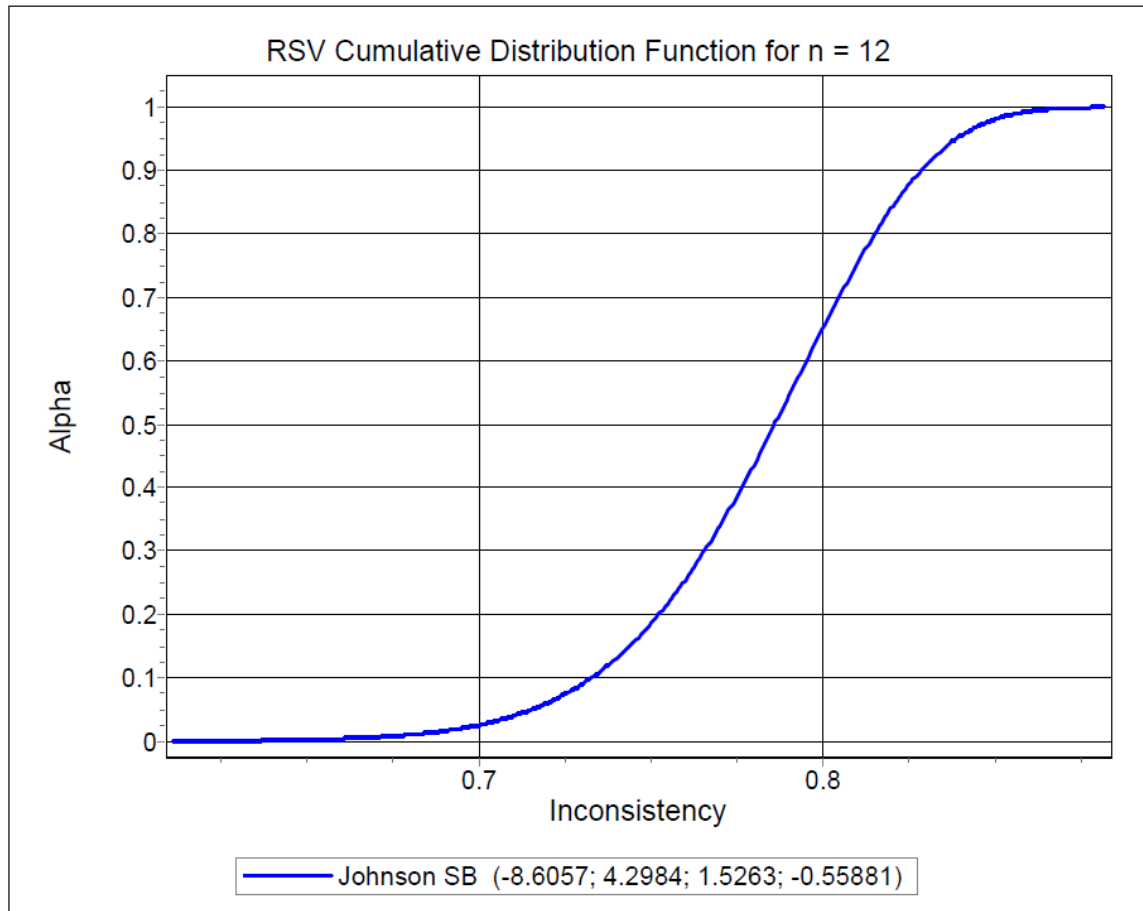


Figure 26: Cumulative Distribution Function (CDF) for 12 Variables