TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-05/0-4393-2		

4. Title and Subtitle	5. Report Date
Implementation of Data Fusion Techniques in Nondestructive Testing of Pavements	October 2004
	6. Performing Organization Code
7. Authors	**8. Performing Organization Report No.**
R. R. Williams, I. Abdallah and S. Nazarian	Research Report 0-4393-2
9. Performing Organization Name and Address	**10. Work Unit No.**
Center for Transportation Infrastructure Systems The University of Texas at El Paso El Paso, Texas 79968-0516	
	11. Contract or Grant No.
	Project No. 0-4393
12. Sponsoring Agency Name and Address	**13. Type of Report and Period Covered**
Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, Texas 78763-5080	Technical Report 9/02– 8/04
	14. Sponsoring Agency Code

15. Supplementary Notes

Research Performed in Cooperation with TxDOT and Federal Highway Administration.
Research Study Title: Integration of Nondestructive Testing Data Analysis Techniques.

16. Abstract

The primary nondestructive testing devices currently in use by TxDOT are the falling weight deflectometer, the seismic pavement analyzer, the portable seismic pavement analyzer, and ground penetrating radar, which provide thickness or modulus information. In many projects a number of these devices are used. Results do not always coincide and thus decisions on either to combine values or decide on one value need to be made. Each device has unique strengths and weaknesses. Data fusion can be used to integrate the results from different devices in a synergistic way by utilizing the strengths of each method while minimizing the weaknesses. Data fusion allows for logical combining and filtering of information to obtain a composite value or a basis for decision. This report presents several data fusion alternatives for the synergistic analysis of data from multiple nondestructive testing devices.

17. Key Words	18. Distribution Statement
Data fusion, Fuzzy logic, Seismic Nondestructive Testing, Deflection Nondestructive Testing	No restrictions. This document is available to the public through the National Technical Information Service, 5285 Port Royal Road, Springfield, Virginia 22161, www.ntis.gov

19. Security Classified (of this report)	20. Security Classified (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	150	

Form DOT F 1700.7 (8-69)

IMPLEMENTATION OF DATA FUSION TECHNIQUES IN NONDESTRUCTRUCTIVE TESTING OF PAVEMENTS

by

Robert Reuben Williams, MSCE
Imad Abdallah, MSCE
and
Soheil Nazarian, Ph.D., P.E.

Research Project 0-4393

Performed in Cooperation with the

Texas Department of Transportation and the Federal Highway Administration

Research Report 0-4393-2
Integration of Nondestructive Testing Data Analysis Techniques

October 2004

The Center for Transportation Infrastructure Systems
The University of Texas at El Paso
El Paso, TX 79968-0516

NOT INTENDED FOR CONSTRUCTION, BIDDING, OR PERMIT PURPOSES

Robert Reuben Williams, MSCE
Imad Abdallah, MSCE
Soheil Nazarian, Ph.D., P.E. (69263)

ACKNOWLEDGMENTS

The authors would like to express their sincere appreciation to the PD of the project John Bilyeu of the TxDOT Construction Division, and German Claros of the Research and Technology Implementation Office for their ever-present support. The project Advisors, consisted of Mike Murphy, Mark McDaniel, Joe Leidy, Dar Hao Chen, Carl Bertrand and Ahmed Eltahan, provided invaluable suggestions to focus the objectives of this project in a manner that is useful to TxDOT day-today use.

We are also grateful to a large number of TxDOT district personnel who assisted us with planning and site selection, and field testing.

We would also like to thank Tom Scullion Co-PI of this project and his staff of researchers for providing us with advice and support on all aspect of this project.

EXECUTIVE SUMMARY

Nondestructive testing (NDT) technology has made substantial progress in the last two decades. Currently, four NDT devices, the Falling Weight Deflectometer (FWD), the Ground Penetrating Radar (GPR), the Seismic Pavement Analyzer (SPA), and the Portable Seismic Property Analyzer (PSPA), are available to TxDOT for collecting field data. Each of these technologies has strengths and weaknesses. However, when combined, they can provide a wealth of information not available when one method is used alone.

The ultimate NDT tool for the evaluation of all pavement systems in Texas would be a device that integrates the capabilities of these NDT tools. The first step toward a fully integrated hardware is a robust integration software. The objective of this project is to harvest the strength of different NDT methods and combine them in a way as to improve the parameters used in pavement design and evaluation. This project will examine the strengths and weaknesses of each device to develop a work plan for integrating information collected from each device in a practical manner.

Developing an algorithm for combining data from different NDT methods with the objective to assess the state of a pavement requires specialized technical capabilities beyond the requirements of conventional data analysis. It requires: (a) a good understanding of each of the NDT techniques being considered, (b) in-depth knowledge of probability and statistical techniques, cross-correlation techniques and techniques for normalizing and re-sampling; and (c) a good understanding of advanced analysis techniques such as artificial neural networks and expert systems. In that context, combining the data from different methods falls under the following two broad categories a) joint inversion, and b) data fusion.

The joint inversion method was described in Research Report 0-4393-1. In this report the data fusion concept is described. Data fusion can be used to integrate the results from different devices in a synergistic way by utilizing the strengths of each method while minimizing the weaknesses. Data fusion allows for logical combining and filtering of information to obtain a composite value or a basis for decision.

IMPLEMENTATION STATEMENT

The primary nondestructive testing devices currently in use by TxDOT are the falling weight deflectometer, the seismic pavement analyzer, the portable seismic pavement analyzer, and ground penetrating radar, which provide thickness or modulus information. In many projects a number of these devices are used. Results do not always coincide and thus decisions on either to combine values or decide on one value need to be made. A user-friendly software package has been developed to implement the data fusion techniques for the devices named above. The software is ready for use on trial basis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

PROBLEM STATEMENT

The most crucial information for assessing the road structural quality is the moduli and thickness of the pavement layers. With several different testing devices available today, there is no shortage of information that can be gathered. However, each method has its own strengths and weaknesses. When the results from multiple testing devices are conflicting, the dilemma arises over which one to accept. Such decisions should be made rationally. The following two approaches can be followed to achieve this goal:

(1) Integrating the input data from different devices into one reduction program. This process is called the Joint Inversion Method (JIM).
(2) Rationally reconciling the results from different methods to arrive at consistent results. This process is called data fusion.

The joint inversion method (JIM) is a backcalculation method that relies on the joint analysis of the raw data from the seismic-based and deflection-based methods. In this type of backcalculation, the inherent strength of each method dominates the analysis, resulting in a more robust and stable algorithm. JIM is not further discussed here since it was thoroughly described in Abdallah et al. (2003) under Research Report 0-4393-1.

The data fusion is a process by which one source of data can be logically selected over another, or by which data from several available sources can be combined or "fused." As each method for analyzing pavements has its own strengths and weaknesses, it is only reasonable to attempt to utilize all methods to develop a better overall characterization of a pavement. As the parameters being measured are not "exact", are subject to inherent errors, all information that has some merit should be considered to some extent.

OBJECTIVE AND APPROACHES

The major objective of this study is to investigate the possibility and feasibility of developing data fusion methods to combine test results from several different nondestructive (NDT) devices. The following three options are considered for data fusion:

1. Statistical/Probabilistic Approaches.
2. Fuzzy Logic Approaches.
3. Hybrid Approaches.

Two main algorithms have been developed for this purpose. In the first algorithm, statistical weighted average method is used. The second algorithm is based on fuzzy logic method. The hybrid approach is a combination of the first two methods. These methods are elaborated further in this report.

ORGANIZATION

The work presented within this report represents the first attempt to use data fusion techniques in the NDT of pavements. Chapter 2 gives an overview of a few of the primary nondestructive test methods that are currently in use, as well as a discussion on the strengths and weaknesses of those methods. The focus of Chapter 3 is on the background on data fusion methods that showed the most promise for this application.

The next four chapters further develop some of the methods discussed in Chapter 2 into four feasible methods for fusing data. Chapter 4 describes a weighted average approach that uses the standard deviation values to develop weights. Chapter 5 presents an adaptation of fuzzy logic such that the statistics obtained from testing and input from experts can be considered. Chapter 6 explains a hybrid method that applies both principles from Chapters 4 and 5 to form a method to combine data. Finally, Chapter 7 introduces an adapted Bayesian approach to make a decision or to fuse data.

Chapter 8 shows applications of the methods to the data from the previous four chapters and to a second set of data. In Chapter 9, conclusions and recommendations for future work on this project is discussed.

CHAPTER TWO

BACKGROUND

INTRODUCTION

In this chapter, nondestructive testing methods used in this study are discussed. The background information behind these methods is briefly described.

NONDESTRUCTIVE TESTING METHODS

Perhaps the biggest concern in testing pavements is to determine the modulus of each pavement layer. Another concern is to accurately measure the thickness of as many layers as possible without having to core the road. Several devices and methods exist for estimating both modulus and thickness. Table 2.1 shows some of the devices currently in use and whether they can be used to determine modulus, thickness, or both. In the remainder of this chapter, the methods shown in Table 2.1 will be discussed.

Table 2.1 - Commonly Used NDT Devices

Device	Modulus	Thickness
Seismic Pavement Analyzer	Yes	Yes
Portable Seismic Pavement Analyzer	Top Layer	Top Layer
Falling Weight Deflectometer	Yes	—
Ground Penetrating Radar	—	Yes

The primary device for collecting structural stiffness data is the Falling Weight Deflectometer (FWD). The Seismic Pavement Analyzer (SPA), the Portable Seismic Pavement Analyzer (PSPA), and the Ground Penetrating Radar (GPR) are often used to complement data obtained by the FWD. The major strengths and weaknesses of these two primary methods are shown in Table 2.2.

Table 2.2 - Major Strengths and Weaknesses of FWD and SPA (after Abdallah et al. 2002)

Method	Strengths	Weaknesses
FWD	Imposes loads that approximate wheel loads	The state-of-stress within pavement strongly depends on moduli of different layers, and hence is unknown
SPA	Measures a fundamentally-correct parameter (i.e., linear elastic modulus)	State-of-stress during seismic tests differs from the state-of-stress under actual loads

Seismic Based Methods

Seismic Pavement Analyzer: The basic principle behind the seismic pavement analyzer (SPA) is the generation and detection of stress waves in a layered medium (Nazarian et al., 1993). The SPA is composed of high and low frequency sources (one of each) and three geophones and five accelerometers to collect data.

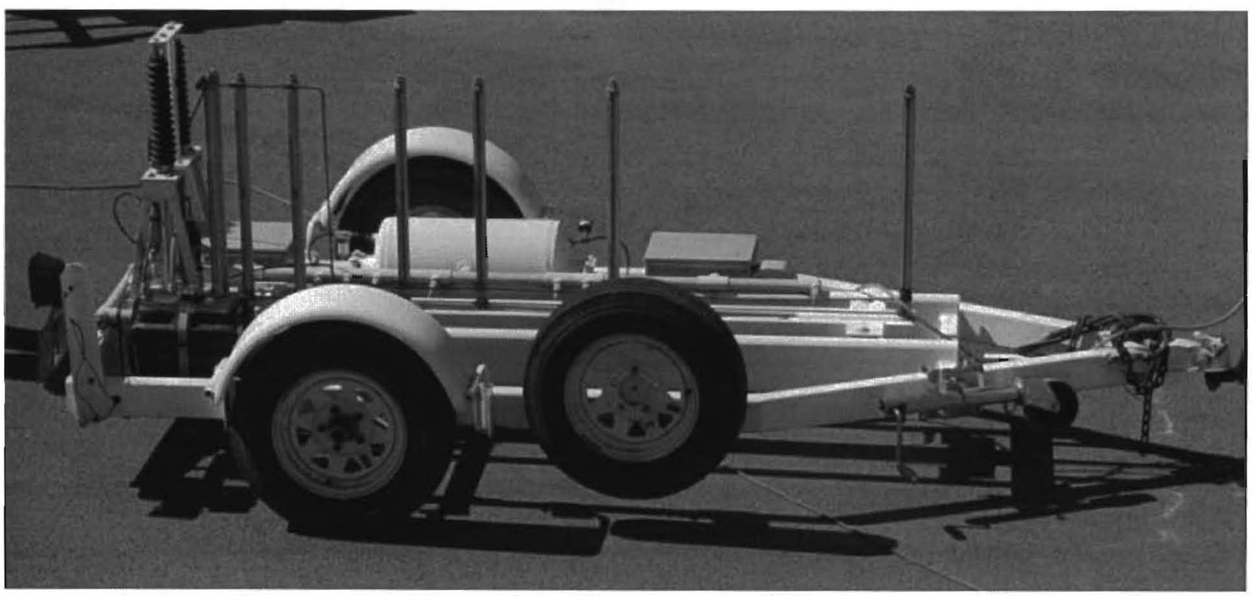


Figure 2.1 - Photograph of SPA

The data collection process consists of the generation of surface waves from the two sources and measuring the motion of the surface with the sensors. The signals are analyzed using Fourier and spectral analysis methods.

The advantages of the SPA include high accuracy in determining the condition of the pavement, that it is based on a sound theoretical background (Nazarian et al., 1993). Additionally, five test techniques are available. These five methods are shown in Table 2.3 along with their primary uses.

Table 2.3 - Methods Available When Using the Seismic Pavement Analyzer (after Nazarian et al., 1993)

Method	Primary Use
Ultrasonic Body Wave	Young's Modulus of top layer
Ultrasonic Surface Wave	Shear Modulus of top layer
Impulse Response	Modulus of subgrade reaction
Impact Echo	Thickness of top layer
Spectral Analysis of Surface Waves	Modulus and Thickness of each layer and variation of modulus in each layer

The method that will be a major focus of this project is the spectral analysis of surface waves (SASW) method.

The impulse-response (IR) method allows for the calculation of two parameters, the shear modulus of the subgrade and the damping ratio of the system. This test requires the use of a low-frequency source and geophone. This test basically measures the response of the entire system.

The ultrasonic surface wave (USW) method is related to the SASW method, but only the top layer is analyzed. Only the high frequency source and the two accelerometers closest to the source are required. This method allows for the computation of the shear modulus of the top layer. If a Poisson's ratio is assumed or known, an estimate of the Young's modulus can then be obtained.

The ultrasonic body wave test is used to obtain the Young's modulus of the top layer. This test merely measures the velocity of the compression wave as it travels through the top layer. As the compression wave is the fastest moving wave, it is the first wave detected on the seismic records.

The impact-echo (IE) method is used to estimate concrete defects. These defects may include voids, cracks, areas of deterioration and the thickness of the top.. The strengths and weaknesses of the first four methods are shown in Table 2.4.

Since much of the focus of this project is on the SASW method, more elaboration will be done for this method. The SASW method does have some limitations and concerns as summarized in Table 2.5. It should be mentioned that the analysis of seismic data is carried out using a software package called Seismic Modulus Analysis and Reduction Tool (SMART) developed by Abdallah et al. (2002) for TxDOT.

In terms of data reduction and analysis, the SASW method consists of two steps: a) the construction of a dispersion curve and b) inversion (backcalculation) of the dispersion curve to obtain a shear wave velocity profile or modulus profile. The two parameters will be used interchangeably since either one of them can be calculated from the other given the Poisson's ratio and density of the material. The dispersion curve is a function that relates frequency to phase velocity. The construction of a dispersion curve is from the phase information of cross power spectra and coherence functions obtained from time records collected from field testing.

Table 2.4 - Advantages and Disadvantages of Four SPA Data Reduction Methods

Method	Advantages	Disadvantages
Ultrasonic Body Wave	• Rapid to perform • Simple data reduction	• Results may be affected by underlying layers • Sensitive to surface condition
Ultrasonic Surface Wave	• Sensitive to properties of top layer • Rapid to perform • Layer specific results	• For multi-course pavements, determination of layer specific information is complex
Impulse Response	• Powerful tool for rapidly locating weak spots in pavement • May be used to estimate depth to stiff layer (under development)	• For flexible pavements, the contribution of different layers are unknown • Results are affected by depth to rigid layer and water table
Impact Echo	• Can determine the thickness of the top layer • Sensitive to delaminated interfaces	• Substantial contrast between the modulus of two adjacent layers is needed for sensitivity • For multi-course pavements, at least one core is needed for calibration • Applies only to pavements with a thicker top layer

Table 2.5 - Advantages and Disadvantages of SASW Reduction Method

Advantages	Disadvantages
Provides the modulus profile in a comprehensive manner	Data reduction is time consuming and complex
More robust than deflection-based methods	Automated analysis applicable only to simple structures

Briefly, a set of phase velocities of surface waves with frequency are calculated, combined and smoothed from several test spacings, and smoothing function is then used to produce a single idealized or representative dispersion curve. Nazarian and Desai (1993) provide a detail account of this process. As an example, a representative dispersion curve obtained from signals received with different spacings is depicted in Figure 2.2. The raw dispersion curves plotted in the graph are a representation of the integrated pavement response under seismic impacts. The ideal dispersion curve is used directly in the inversion process to estimate pavement properties.

The second part of the SASW method, the inversion process, is rather complex. In the inversion process the dispersion curve is assumed as the input data. The complexity in the process mainly depends on the amount of information that is being estimated. Theoretically, the information contained in the dispersion curve includes both thicknesses and shear wave velocities of the pavement layers. However, the more unknown variables that are estimated, the more uncertainty exist in the analysis.

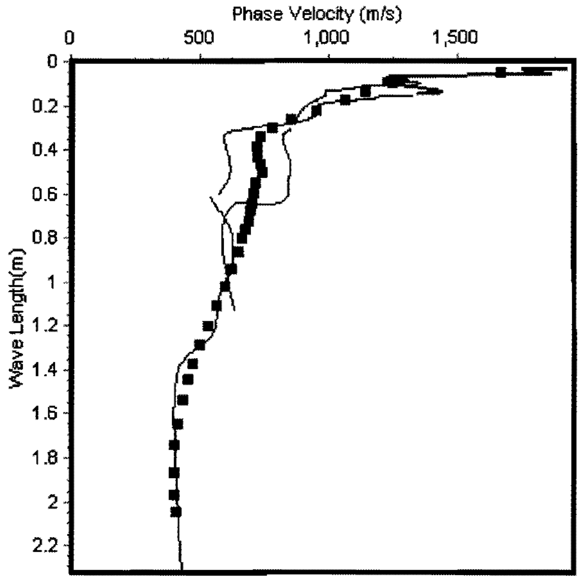


Figure 2.2 - Idealized Dispersion Curve

As stated, inversion is an iterative process in which a theoretical dispersion curve is calculated for an assumed shear wave velocity profile. The theoretical and ideal dispersion curves are compared. Based on the errors between the two curves, the assumed profile is modified to improve the match. This process is continued until error or difference between the two curves is minimized. Yuan and Nazarian (1993) describe the automation process for all three steps. Figure 2.3 shows graphically how the inversion process is carried out. The first, third and seventh iterations are presented as an example of conversion by the algorithm. In this example the profile from the seventh iteration produces theoretical dispersion curve that provides a good fit to the experimental dispersion curve. The profile is then used as an estimation of the material properties.

The overall SASW reduction algorithm is illustrated in Fig. 2.4. The first step shows the input required. The first input is the raw data. This information is contained in several files representing information from different sensor spacings. The other set of information is a priori information, the best guess about the properties of the system. The a priori information includes the number of layers, layer thickness, and the type and quality of different layers. Based on the type and quality, a set of modulus values (shear wave velocity values) are selected. These values can always be modified to suit the system being evaluated. In Step 2 the raw data is extracted for each sensor spacing and phase spectra are calculated and plotted. In Step 3, phase spectrum from each spacing is unfolded and used to develop a dispersion curve. The final part in Step 3 is a curve fitting algorithm that produces an ideal dispersion curve that represents the raw data. In Step 4, the inversion process, the idealized dispersion curve and a priori information are input and modified to match the idealized dispersion curve with the theoretical dispersion curve. The stiffness profile that produces the closest match between the theoretical and experimental dispersion curves is selected as the final output profile (Step 5). The latest versions and improvements of the algorithm including both the dispersion and inversion process are discussed next.

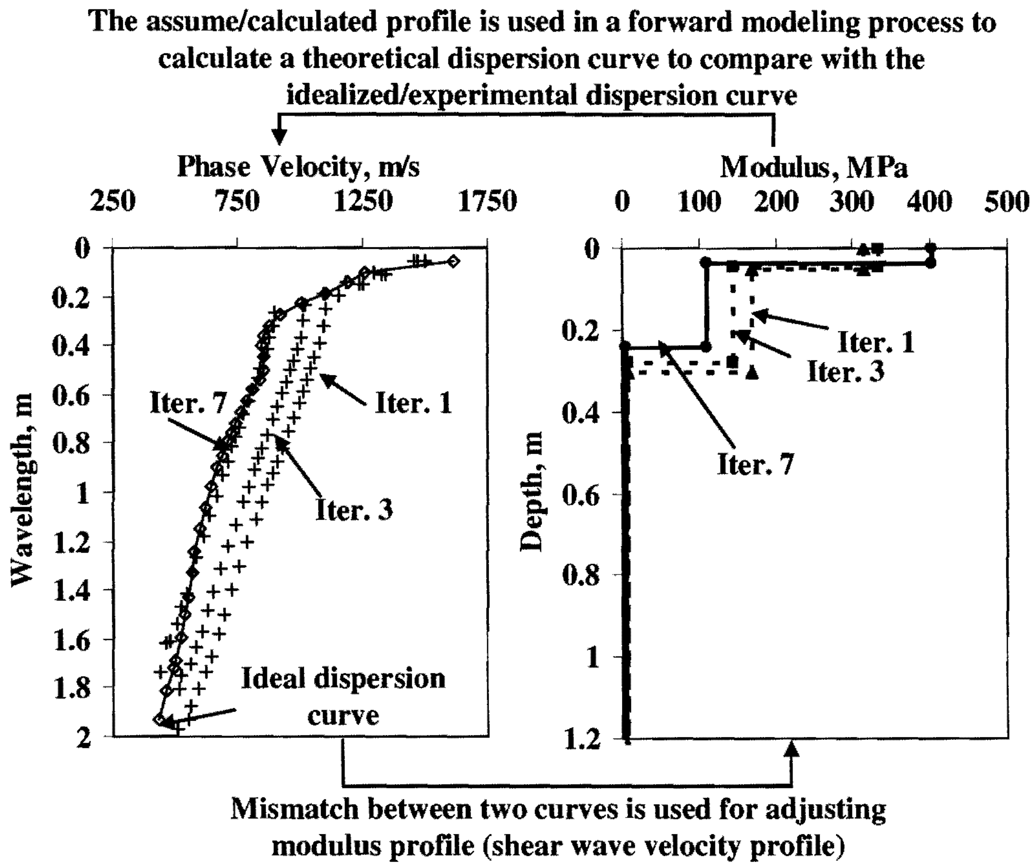


Figure 2.3 - Graphical Illustration of SASW Inversion Process

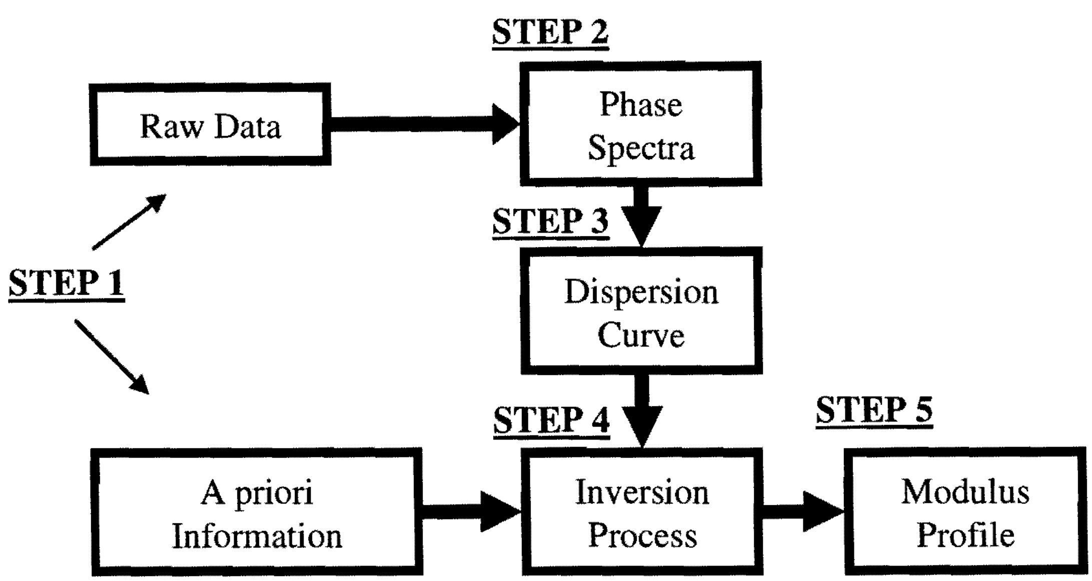


Figure 2.4 - Data Reduction and Analysis Procedure

The final step in the process is to determine the design moduli. A schematic of the algorithm used by SMART to provide the design modulus for each layer is shown in Figure 2.5. After the SASW analysis is completed, the moduli are fit into appropriate material models to obtain moduli that would be experienced under wheel load (see Abdallah et al., 2002 for details).

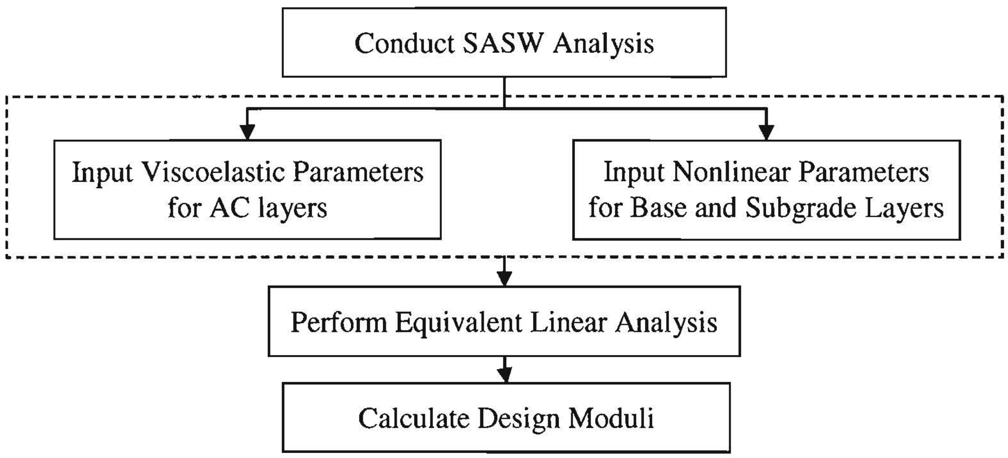


Figure 2.5 - Flowchart Illustrating Process in SMART for Determining Seismic Modulus

Portable Seismic Pavement Analyzer: The portable seismic pavement analyzer (PSPA) can be thought of as a smaller version of the SPA that is used for testing the top pavement layer only. Unlike the SPA, the PSPA only uses a high-frequency source and two accelerometers that serve as the receivers. Figure 2.6 shows a photograph of the PSPA.

The PSPA data is primarily reduced using the USW method or the IE method using SPAManager software. The USW method used in PSPA data reduction only considers that the waves are contained within the top layer. As such, the velocity measured is the velocity of the top layer only. The modulus of the layer can then be calculated from the velocity. If the layers are thick enough (as in most concrete or with thick asphalt) the thickness of the layer can be found using the IE method.

Figure 2.6 - Photograph of PSPA

Falling Weight Deflectometer

The falling weight deflectometer (FWD) is a deflection based device that operates on the principle of applying an impulse load to a pavement and then recording the surface deflections at predetermined intervals. The current setup uses seven geophones spaced at one foot intervals for a total span of six feet (the first geophone is positioned at the center of the load plate).

Raw data from the FWD test consist of the applied load and resulting seven deflections. The sensors farther from the source will provide information about the deeper layers (subgrade) and the closer sensors provide information about the upper layers. Figure 2.7 shows a typical FWD.

Figure 2.7 - Photograph of FWD

Data from the FWD is primarily reduced by minimizing the errors between the measured deflections to those theoretically determined for an assumed pavement. Uzan et al. (1988) describe the principles behind the backcalculation procedure as used in the MODULUS program. MODULUS uses a linear-elastic algorithm to generate a database of deflection bowls for assumed profiles. The idea is to find a theoretical deflection bowl that matches that of the measured deflection bowl. A search procedure to find the closest fitting deflection bowl (minimum error) is used. The number of generated deflection bowls depends on the range of modulus values entered by the user.

The error that is minimized is a function of the measured and estimated deflections and a weight factor for each deflection. A Hooke-Jeeves pattern recognition search algorithm is used, as it always converges although occasionally to a local minimum. The pavement profile that is selected is the profile with the smallest error.

The MODULUS software is simple to use and requires little training. The strengths and weaknesses for this method are reiterated in Table 2.6. The pavement thicknesses must be known *a priori*. If the thickness of the top layer is less than three inches, the modulus should be assumed.

Ground Penetrating Radar

The ground penetrating radar (GPR) is a fast and efficient way to estimate the thickness of pavement layers. The thickness of the top (usually asphalt concrete) layer is obtained, as well as the thickness of the lower layers. Ground penetrating radar can also be used to find the dielectric properties of the layers, a parameter that can give some indication of the moisture content of the layers.

Table 2.6 - Advantages and Disadvantages of MODULUS Program for FWD Data

Advantages	Disadvantages
Easy to use	Not sensitive to thin pavements
	Increased error/variability caused by shallow bedrock
Fast reduction	No way to backcalculate thickness
	Increase in variability as the number of layers modeled increases
	Subgrade (bottom) layer moduli may not be as constant as backcalculated results would indicate
Already widely used	Results and error can vary depending on the range of initial model
	Impossible to competently reduce without reliable layer thickness information

GPR operates on the principle of measuring reflected electromagnetic energy that bounces off the soil layer interfaces after being transmitted from an antenna in short pulses (Maser et al., 1991). The differences in the layers are found by measuring the dielectric discontinuities that are encountered when the signal moves from one layer to another. The arrival time and amplitude of the reflected signals are important when using GPR for determining the dielectric values and the layer thicknesses. The thicknesses are calculated from the arrival times and amplitudes of the returning signals and the dielectric constant are dependent on the amplitudes of the bounced back signals (Maser et al., 1991). The accuracy of GPR in determining thicknesses has been up to ±5-7.5% for asphalt and ±9-12% for base (Maser and Scullion, 1992).

Figure 2.8 shows the GPR device. It is simply a van housing data acquisition hardware/software, with an antenna mounted on a boom off the front bumper. The antenna is suspended approximately 10 to 14 inches above the ground. The system can operate easily over open road. Scullion et al. (1994) describe the details of the system hardware and data reduction.

Figure 2.8 - Photograph of GPR

Ground penetrating radar operates at highway speeds. This is a major benefit because it allows for a long stretch of roadway to be tested in a relatively short period of time. Additionally, GPR takes a reading at set intervals which allows for a near-continuous thickness profile of the pavement to be obtained.

Since GPR depends highly on the dielectric properties of the pavement layers, it is sensitive to moisture. This has both an upside and a downside. The upside is that if a material has become inundated with water that will be evident when the GPR test is performed. However, if the dielectric properties of the base and underlying layer become too similar due to the presence of water, then the test may not be able to yield a distinction between those layers.

The ground penetrating radar is a fast and easy way that can obtain a near-continuous profile for the thickness of pavement layers. Using GPR in concert with other methods can help to obtain better overall information about the site.

CHAPTER THREE

DATA FUSION METHODS

In any pavement evaluation process, accuracy and precision are important. When a parameter is measured, the resulting measurement contains error. The ideal situation would be the elimination or reduction of errors. However, that may prove impossible without extraordinary effort. Where multiple devices are used to measure the same parameter, the devices can either support or contradict each other. Data fusion is a general term used to describe the processes by which results from different devices are combined with the purpose of arriving at more reliable or accurate results than the individual devices would arrive on their own. These readings can be both reduced results and raw data, depending on the level data fusion that is being applied. The complexity of data fusion can range from a simple average to an artificial neural network (Gros, 1997). Numerous data fusion methods are in use today and the number continues to increase as current methods change and new methods are introduced.

CONCEPT OF DATA FUSION

One of the main bases for data fusion is that any event has some degree of "truth" associated with it. This truth can be represented by a probability, by a weight, by a distribution, or by a plausibility interval. It is usually possible to obtain a weight from any of the methods just listed. Weights, in turn, can be used to combine readings from multiple devices. The fundamental challenge to any data fusion method is how to apply the method to obtain weights which can then be used to arrive at a composite value. Furthermore these weights should describe a particular event or should identify potential outliers so that they can be adjusted or eliminated. Some methods are more conducive to manipulating and modifying data, whereas others only lend themselves to identifying potentially troublesome data points. Once a systematic method for data manipulation and data filtering is defined, algorithms can then be developed to remove the burden of guess work for a person trying to make a decision.

A conceptual example of data fusion is how to rate a restaurant. A person will take in information from all of his five senses, as shown in Figure 3.1. To rate the food, the person will smell and taste it. Additionally, pleasant or unpleasant ambient odors may impact his perception of the restaurant. Also, the person will hear and see things that will help him determine the atmosphere of the restaurant. "How noisy is it?" or "Is the décor pleasant to look at?" will be factors that he notices that impact his impression of the restaurant. The fifth sense, touch, will give the patron an idea of how comfortable the restaurant is. For example it may be too warm or too cold, there may be an unpleasant draft, or the seating may be uncomfortable. In addition to the senses, other factors may impact the rating of the restaurant such as the quality of service or the reasonableness of the price of the food. So in addition to the five senses, there are two other factors that can also be considered in evaluating the restaurant. Exceptional food may outweigh high prices, while bad service may outweigh excellent food. The fusion that will take place in the restaurant patron's mind will take into consideration all of those factors to come to a final perception of the restaurant.

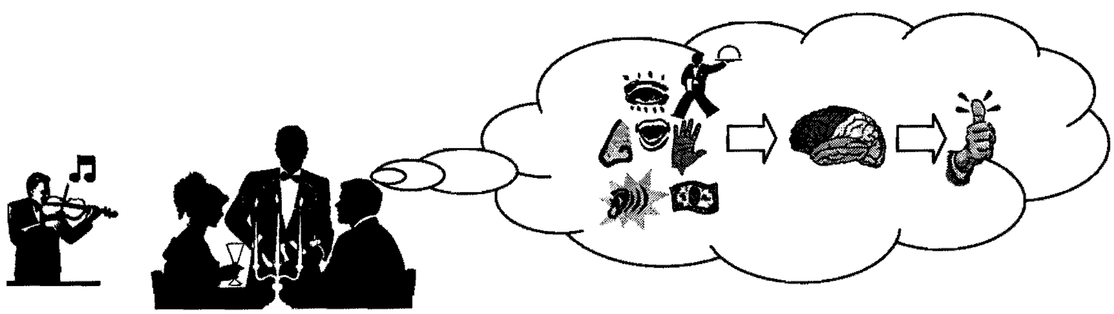


Figure 3.1 - Concept of Data Fusion

Data fusion of experimental data is very similar, there may be data from several sources, each with its own advantages and disadvantages, and a more complete view of the phenomenon or the parameter being measured can be obtained from the systematic synthesis of this data.

DATA FUSION CATEGORIES

Data fusion methods can be separated into different categories depending on the main characteristics of the method. Data fusion methods have been broken up into three categories. The first category is statistical and probabilistic approaches to data fusion. The methods contained within are a weighted average that relies on statistical information such as mean and standard deviation. The other method is the Bayesian inference method which requires the use of prior and posterior probabilities in determining the likelihood of an event (Gros, 1997). The next category of data fusion is evidential reasoning. The Dempster-Shafer evidential reasoning method allows for expert opinions from different sources to be combined based on where these different opinions intersect (Gros, 1997; DeCETI, 2000). The Dempster-Shafer method returns a result of a confidence interval, or range of probabilities for which a given event can occur. The third category is fuzzy logic. The fuzzy logic method allows for the creation of rules that attempt to function in a manner that changes a subjective decision into a mathematical rule. These decisions, in turn, can be used to combine or filter data.

Statistical/Probabilistic Approaches

Statistical Weighted Average: Most measurements have a level of uncertainty associated with them. Thus any method that takes that into consideration would be desirable. One main limitation is the difficulty in calculating that uncertainty. The statistical weighted average uses simple statistical values to combine data. From most sets of numbers, a mean and standard deviation can be obtained. Additionally, a standard deviation can be assumed in cases where one cannot be calculated. This usually happens when the size of the sample is small.

Taylor (1997) shows the development of a method based on the Gaussian distribution. Consider several Gaussian distributions that can be described by the following probability for a value X that maximizes the probability

$$P_X(x_i) \propto \frac{1}{\sigma_i} e^{-(x_i - X)^2/2\sigma_i^2} \quad \text{for i = 1, 2,..., n} \tag{3.1}$$

For each i, the probability depends on the exponent term as the probabilities depend on X. The product of all of the probabilities can be represented as

$$P_X(x_1, x_2 ..., x_n) = \prod_{i=1}^{n} P_X(x_i) \propto \prod_{i=1}^{n} \frac{1}{\sigma_i} e^{-\chi^2/2} \tag{3.2}$$

where the variable χ^2 is given by

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i - X}{\sigma_i} \right)^2 \tag{3.3}$$

Since the probability will be greatest for each i where $x_i - X$ is a minimum, the first derivative test can be applied to find the X for which the χ^2 function will be minimal, hence

$$\sum_{i=1}^{n} 2 \frac{x_i - X}{\sigma_i^2} = 0 \tag{3.4}$$

The solution to the previous equation is

$$X = \frac{\left(\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \right)}{\left(\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \right)} \tag{3.5}$$

By extension, the individual weights can be defined as

$$w_i = \frac{1}{\sigma_i^2} \tag{3.6}$$

the combined value of x will be

$$x_F = \frac{\sum w_i x_i}{\sum w_i} \tag{3.7}$$

From the principle of error propagation the combined uncertainty will be

$$\sigma_F = \frac{1}{\sqrt{\sum w_i}} \tag{3.8}$$

To derive the composite uncertainty function (Equation 3.8) the use of the principle of error propagation is required. Two rules of error propagation must be applied, one that accounts for scalar multiplication and one that accounts for a summation of variables. The rules are as follows in respective order.

$$f(x) = Kx \Rightarrow \sigma_{f(x)} = |K| \sigma_x \tag{3.9}$$

and

$$f(x_1, x_2, ..., x_n) = x_1 \pm x_2 \pm ... \pm x_n \tag{3.10a}$$

$$\sigma_{f(x_1, x_2, ..., x_n)} = \sqrt{(\sigma_{x_1})^2 + (\sigma_{x_2})^2 + ... + (\sigma_{x_n})^2} \tag{3.10b}$$

Substituting Equation 3.7 into Equation 3.10 and applying the rule for constants from Equation 3.9, the following equation is derived

$$\sigma_F = \sqrt{\frac{\sigma_1^2 w_1^2 + \sigma_2^2 w_2^2 + ... + \sigma_n^2 w_n^2}{(w_1 + w_2 + ... + w_n)^2}} \tag{3.11}$$

Since $w_i = \frac{1}{\sigma_i^2}$, the preceding equation simplifies to

$$\sigma_F = \sqrt{\frac{w_1 + w_2 + ... + w_n}{(w_1 + w_2 + ... + w_n)^2}} = \frac{1}{\sqrt{\sum w_i}} \tag{3.12}$$

Let us consider an example with four values, each with their own mean and standard deviation. The values for the distributions are shown in Table 3.1 and the plot of the probability density function is shown in Figure 3.2.

A key feature of this method is taking into consideration not only a given value, but its uncertainty as well. With the traditional method, which is just a regular mean and standard deviation, the prior knowledge of each values uncertainty is completely ignored. However, in the weighted method, those uncertainties are taken into consideration. Note how the uncertainty of the weighted result is lower than that of the traditional result. The whole reason behind using this weighted method is not only to weight more reliable measurements more heavily, but also to develop a composite uncertainty.

Table 3.1 - Statistical Weighted Average Example

Case	x_i	σ_i	w_i	$w_i x_i$
1	56	14	$^1/_{196}$	$56 \times {}^1/_{196}$
2	36	12	$^1/_{144}$	$36 \times {}^1/_{144}$
3	38	9	$^1/_{81}$	$38 \times {}^1/_{81}$
4	44	12	$^1/_{144}$	$44 \times {}^1/_{144}$
		Σ	0.0313	1.31
	Traditional		**Weighted**	
	$\mu =$ 43.5		$x_F =$ 41.8	
	$\sigma =$ 9		$\sigma_F =$ 5.6	

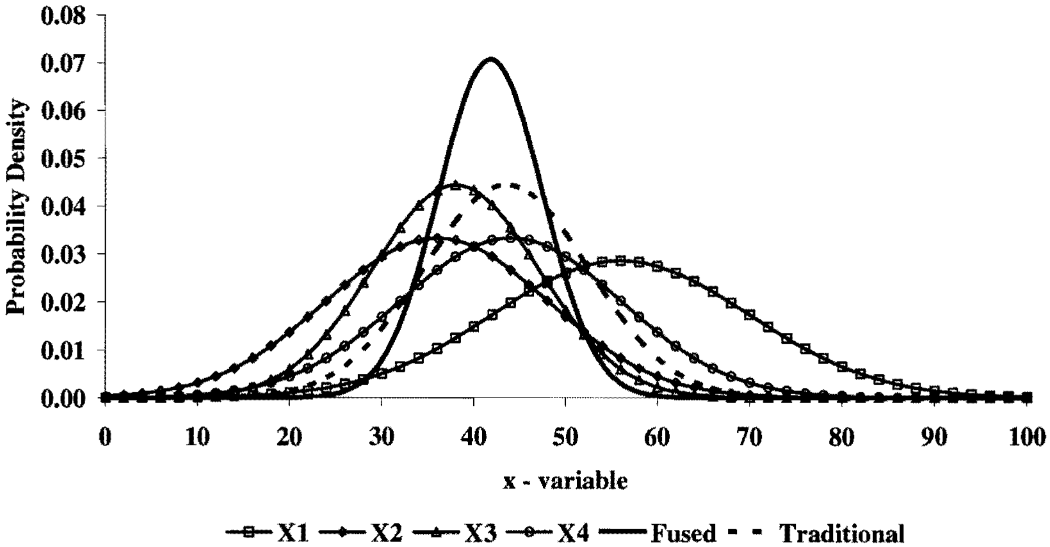


Figure 3.2 - Probability Density Functions for Statistical Weighted Average Example

The strength of this method is that it can be applied to any number of data points with a simple calculation. Once a mean and standard deviation are known for each data point, the numbers can easily be combined. Ang and Tang (1975) proposed a similar process to combine data, but only for two variables.

Bayesian Inference: The role that Bayesian inference plays in data fusion is eliminating data that may be unlikely and eliminate ambiguities and conflicting information (Gros, 1997). The result of using Bayesian inference would be the selection of one device data over another. However, this can be extended to use the resulting probabilities to create weights that could later be used to fuse results from multiple sensors.

The Laboratory of Remote Sensing (DeCETI, 2000) described a method for applying Bayesian inference to a distributed variable. This method relies on *a priori* probabilities and conditional density functions, that is $P(E_i)$ and $p(x|E_i)$, respectively. The probability density function can be represented by any distribution. Normal distributions will be used to illustrate how to apply the Bayesian method to a continuous variable. The *a priori* probability will be converted to an *a posteriori* probability, $P(E_i|x)$. Once the *a posteriori* probabilities have been found, a decision on which input to select can be made. The posterior probabilities of the events are defined by taking the relative ratio of the weights for each input to the total input as given by the Equation 3.13. The denominator in the equation is a normalization factor so that the sum the probabilities of all of the events being compared is one.

$$P(E_j \mid x) = \frac{p(x \mid E_j) \cdot P(E_j)}{p(x)} \tag{3.13}$$

$$p(x) = \sum_{j=1}^{n} p(x \mid E_j) \cdot P(E_j) \tag{3.14}$$

An example can be introduced at this point to illustrate how the Bayesian inference method can be used to decide among three options for the value of a given event. For argument's sake, let the phenomenon being reported be gas mileage for a car. Three different experiments were conducted, each with their own mean and standard deviation. Based on the given data, the gas mileage has to be determined. Assuming a Gaussian distribution for each experiment, the probability density functions and the probabilities of each even occurring can be found, and a decision can be made. The data for this example is shown in Table 3.2.

Table 3.2 - Distribution Data for Bayesian Inference Example

Experiment	Average (*mpg*)	Standard Deviation (*mpg*)
E_1	28	4.2
E_2	22	2.2
E_3	34	6.8

Now let us assume that the measurements that were made to arrive at the distributions in Table 3.1 are looked at individually. Furthermore, let us assume that the final reading from each set of experiments will be considered. There are two options for calculating the probability of a given value based on its distribution. One is the cumulative probability approach and the other is the double-tailed probability. Figure 3.3 shows the difference between the two options for finding the probability. The areas under the black curves are what are included in the probability calculation whereas the gray curves show the entire distributions.

Table 3.3 shows the results of applying Equations 3.13 and 3.14 to the data shown in the second column of Table 3.2 using the distributions from Table 3.1.

From Table 3.3, the result that would be selected based on the highest posterior probability would be E_1. Another option would be to use the posterior probabilities as weights for fusing the data. If a weighted average is taken using the posterior probabilities as weights, the fused result is 27.6 mpg with a 2.5 mpg standard deviation. The uncertainty is derived in using the principles of error propagation that were described earlier. The equation used for the uncertainties is

$$\sigma_F = \sqrt{\frac{\sum (w_i \sigma_i)^2}{\left(\sum (w_i)\right)^2}} \tag{3.15}$$

where the weights are the posterior probabilities and the standard deviations are the standard deviations for each experiment.

The Bayesian inference method that was proposed by the Laboratory for Remote Sensing relies on being able to establish the conditional density functions and the *a priori* probabilities (DeCETI, 2000). It shows how those two pieces of information can be arrived at fairly easily. The rest of the procedure simply relies on performing the necessary calculations using Equations 3.13 and 3.14.

Table 3.3 - Summary of Bayesian Prior and Posterior Probabilities

Distribution	Value	Prior Probability		Posterior Probability	
		Cumulative	Double-tailed	Cumulative	Double-tailed
E_1	28	0.50	1.00	0.50	0.76
E_2	26	0.97	0.07	0.35	0.02
E_3	30	0.28	0.56	0.14	0.22

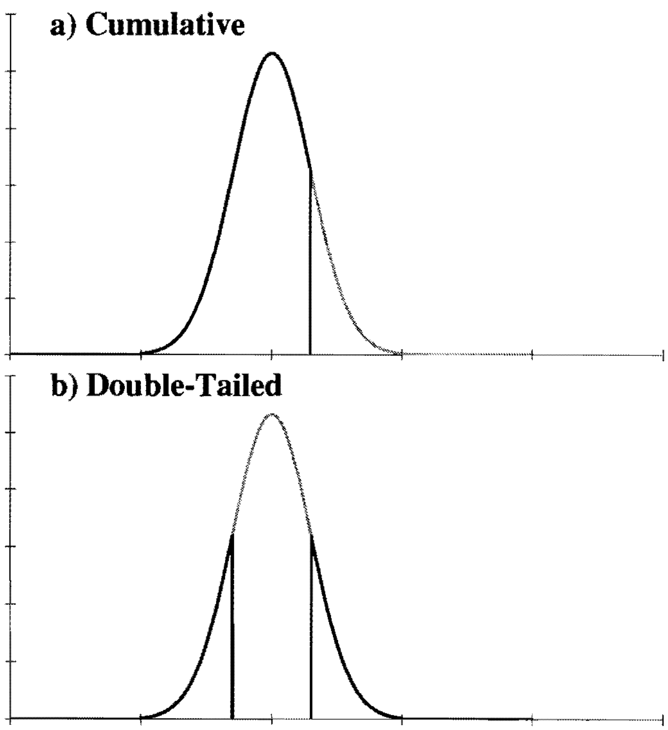


Figure 3.3 - Options for Probability Calculation

The Bayesian method is an excellent way to compare data from several sources at once. In the case of this example, the prior probabilities were found using a normal distribution. However, if data is not available to find a prior probability and there are multiple options, the prior probability of each can be assumed to be equal, $^1/_N$, where N is the number of experiments. However, if all of the prior probabilities have a value of one the result will be the same because the probabilities are normalized as Equation 3.14 is a normalizing factor. If the conditional density functions are not known, conditional probability values can be substituted instead. Generally, those conditional probabilities are derived from empirical records, computer simulations such as artificial neural networks, or they can be assumed outright or with an assigned distribution.

Bayesian inference allows for the comparison of many options with calculations that are not too difficult to perform. The only problem that arises is that the input for the equations can be difficult to develop. If an algorithm for developing the inputs is available, this method can be very powerful.

Evidential Reasoning Approaches

Dempster-Shafer Method: The Dempster-Shafer method derives conclusions from combining evidence (Gros, 1997; DeCETI, 2000). This method relies on combining sets of data that either partially or fully support each other. An easy way to conceptualize this is to suppose there are three friends and one of them breaks a neighbor's window. One witness thinks it was person one or two, another thinks that it was person one or three. Intuitively, the inclination is to believe that person one is the culprit. This is the main premise behind the Dempster-Shafer method, only more sets with multiple subsets can be combined. The Dempster-Shafer method then takes these intersecting events and develops a belief interval that has an upper and lower probability of the event occurring.

These upper and lower probabilities are referred to as belief and plausibility. There is a difference between belief and plausibility. Belief is the minimum probability of an event occurring and plausibility the maximum probability. This means that the plausibility of an event must always be equal to or greater than the belief. Figure 3.4 shows an example of a confidence interval

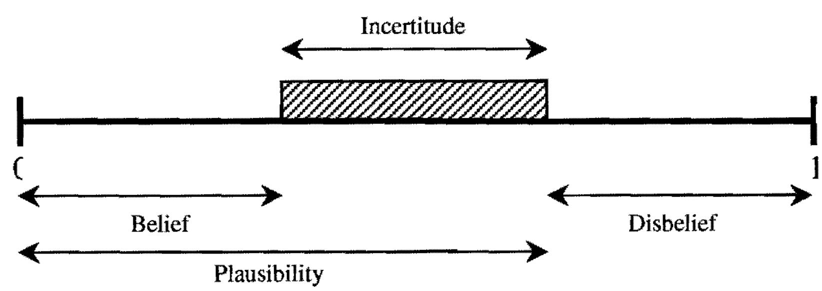


Figure 3.4 - Confidence Interval

The Dempster-Shafer method is based on the theory of evidence (Gros, 1997) to assign a belief or probability to a given event. Belief functions are used to represent the degree of faith in a possible outcome. The belief function of the initial data is basically a probability assignment. This probability can be assigned to data by using a probability distribution, statistical information, or judgment. For now it will be assumed that all probabilities have been appropriately assigned. A belief function, Bel(x), where a proposition has multiple hypotheses can be denoted as

$$Bel(X) = \sum_{X_i \subseteq X} m(X) \tag{3.16}$$

The X_i's refer to the hypotheses that are contained in proposition X. Each event X_i contained in X can be assigned its own probability. The probability of each hypothesis can be referred to as $m(X_i)$. The next part of the procedure calls for Dempster's rule of combination. This can be expressed mathematically as

$$m_1 \oplus m_2 = m_3(Z) = K \sum_{X_1 \cap X_2 = Z} m_1(X_1) m_2(X_2) \tag{3.17}$$

where K is a normalization factor that will be derived shortly. This rule uses the orthogonal sum of sets that are to be combined. This equation can be rewritten as

$$Bel(Z) = K \sum_{\substack{i,j \\ X_1 \cap X_2 = Z}} m_1(X_i) m_2(X_j) \tag{3.18}$$

Equation 3.18 basically multiplies any two hypotheses (one from each set). The new set formed will be referred to by the name of what the two combined sets have in common. If two hypotheses do not support each other, i.e. have nothing in common, then a value, k, can be found which is used to calculate the normalization factor.

$$k = \sum_{\substack{i,j \\ X_1 \cap X_2 = \emptyset}} m_1(X_i) m_2(X_j) \tag{3.19}$$

The normalization factor is simply

$$K = \frac{1}{1-k} \tag{3.20}$$

The final concept to be introduced is the plausibility of an event happening. Plausibility does not consider what two hypotheses have in common, but where they overlap. This can be expressed as

$$Pls(Z) = 1 - Bel(\overline{Z}) \tag{3.21}$$

The plausibility is the difference between total belief and the belief that an event will definitely not occur.

The application of these equations and their use in the Dempster-Shafer method can best be illustrated with an example. Suppose that one was trying to determine who did break the window, Larry, Moe, or Curly. There are three possibilities for who the culprit is, but from gathered evidence there are several possibilities which can overlap, creating other members of the set besides the three simple cases. Some witnesses may only report seeing a man while another may say that the suspect had hair, thus creating a possibility of more than one suspect from one observation. Suppose that information is gathered from two investigators and is as shown in Table 3.4. The two investigators will be referred to as Investigator X and Investigator Y, or more simply as sets X and Y. Beliefs and plausibilities can be established for each parameter before the two sources are combined these values are also shown in Table 3.4.

Table 3.4 - Initial Probabilities, Beliefs, and Plausibilities

Combination	P(X)	Bel(X)	Pls(X)	P(Y)	Bel(Y)	Pls(Y)
L*	0.3	0.3	0.9	0.4	0.4	1
M*	0	0	0.7	0	0	0.5
C*	0	0	0.5	0	0	0.6
L, M	0.2	0.5	1	0	0.4	1
L, C	0	0.3	1	0.1	0.5	1
M, C	0.1	0.1	0.7	0	0	0.6
L, M, C	0.4	1	1	0.5	1	1

* L=Larry, M=Moe and C=Curly

The belief for set X with Larry or Moe would be $Bel(X_{L,M}) = P(X_L) + P(X_M) + P(X_{L,M})$. The plausibility for the same situation would be $Pls(X_{L,M}) = P(X_L) + P(X_M) + P(X_{L,M}) + P(X_{L,C}) + P(X_{M,C}) + P(X_{L,M,C})$ or $Pls(X_{L,M}) = 1 - P(X_C)$.

Now the two inputs must be combined by multiplying the two sets of masses together The outputs that each will support after multiplication will be what the two masses have in common. If they have nothing in common, then the result is part of the empty set. The two inputs X and Y will be broken into four and three masses, respectively. These will be denoted as x_i and y_j. Table 3.5 below shows the masses.

Table 3.5 - Masses for Multiplication

Set X	Set Y
$x_1\{L\} = 0.3$	$y_1\{L\} = 0.4$
$x_2\{L, M\} = 0.2$	$y_2\{L, C\} = 0.1$
$x_3\{M, C\} = 0.1$	$y_3\{L, M, C\} = 0.5$
$x_4\{L, M, C\} = 0.4$	

The product of multiplying sets X and Y will be Z and the masses for Z will be referred to as z_{ij}, e.g. $x_3\{M, C\} \times y_2\{L, C\} = z_{32}\{C\}$. The result of multiplying the masses will be twelve new masses, which can then be combined using evidential reasoning. Table 3.6 shows the twelve new masses.

Table 3.6 - Composite Masses

Set Z	
$z_{11}\{L\} = (0.3)(0.4) = 0.12$	$z_{31}\{\emptyset\} = (0.1)(0.4) = 0.04$
$z_{12}\{L\} = (0.3)(0.1) = 0.03$	$z_{32}\{C\} = (0.1)(0.1) = 0.01$
$z_{13}\{L\} = (0.3)(0.5) = 0.15$	$z_{33}\{M, C\} = (0.1)(0.5) = 0.05$
$z_{21}\{L\} = (0.2)(0.4) = 0.08$	$z_{41}\{L\} = (0.4)(0.4) = 0.16$
$z_{22}\{L\} = (0.2)(0.1) = 0.02$	$z_{42}\{L, C\} = (0.4)(0.1) = 0.04$
$z_{23}\{L, M\} = (0.2)(0.5) = 0.10$	$z_{43}\{L, M, C\} = (0.4)(0.5) = 0.20$

The next step will be to combine like masses and multiply by the K factor ($1/(1 - k)$, $k = z(\emptyset)$). The multiplication by the K factor is to normalize the usable results to one so that the beliefs and plausibilities can be found. Table 3.7 shows the results of these procedures. The *P(Z)/K* column is the non-normalized result that omits the empty set data from z_{31}; the *P(Z)* is normalized and the probabilities add up to one.

Table 3.7 - Results for Combined Evidence

Combination	P(Z)/K	P(Z)	Bel(Z)	Pls(Z)
L	0.56	0.58	0.58	0.94
M	0	0.00	0.00	0.36
C	0.01	0.01	0.01	0.31
L, M	0.1	0.10	0.69	0.99
L, C	0.04	0.04	0.64	1.00
M, C	0.05	0.05	0.06	0.42
L, M, C	0.2	0.21	1.00	1.00

From the results, the most likely suspect for the window breaking can be determined based on the combined evidence. The evidence refutes that the guilty party is either Moe or Curly, [0, 0.36] and [0.01, 0.31], respectively ([*Bel, Pls*]). Both of the limits of the two belief intervals are below the one-half point, thus refuting a belief in Curly or Moe being responsible. The evidence does support Larry, Larry-Moe, and Larry-Curly, but knowing that the evidence against Moe and Curly is refragable places doubt on either Larry-Moe or Larry-Curly being a reasonable choice. By process of elimination, it can be said that the person most likely to have broken the window is Larry.

If there was a third investigator, say another set called W, then the data could have been combined in any order since the Dempster-Shafer method is both commutative and associative. (Gros, 1997).

The Dempster-Shafer method is an effective way to combine sets that overlap each other in a way to determine the most likely outcome. When the Dempster-Shafer method is broken up into non-overlapping sets and applied to Gaussian distributions, the results are identical to the weighted average method once proper scaling factors are applied. However, it is much simpler to just apply the weighted average method than to use Dempster-Shafer to fuse Gaussian distributions.

Fuzzy Logic

Fuzzy logic was first proposed by Zadeh (1965). Fuzzy logic is an extension of conventional set theory, binary logic, and probability measure (Tanaka 1991). Fuzzy logic tries to handle vagueness or ambiguity that is the result of human thinking (Tanaka 1991). The idea behind fuzzy logic is a way to imitate the decision process that relies on making a subjective decision with a decision algorithm formed by defining rules that will interpret the inputs and make a decision based on those rules (Jantzen, 1997; Jantzen, 1998; Favata, 2001; Mathworks, 2003). Input for fuzzy logic is not necessarily, and usually not, discrete, (i.e. true or false, yes or no, on or off). Fuzzy logic allows for a gray area in between black and white.

Fuzzy logic depends on the development of fuzzy sets. These fuzzy sets have no definite boundary. Events that occur within these sets often partially belong to two members of the set. The following example illustrates the concept of fuzzy sets.

A baseball player is considered to be a power hitter if he has more than 35 home runs and an average hitter if he has 20 home runs. What about a player with 27 home runs? All that can be ascertained about the hitter is that he is an above average hitter with more than 20 home runs. The fact that more than 35 home runs would qualify him as a good hitter has no influence on our reasoning if we use Boolean logic. Fuzzy logic allows for these situations to be dealt with by representing a degree of membership in each of the categories.

Now the concept of a membership function must be introduced. In Boolean logic, the value 27 home runs would constitute the hitter as not a good hitter. However, intuition tells us that this hitter is still a decent player. Figure 3.5 shows that everyone below 35 is not to be taken as a good hitter. The membership function only has two options, 0 or 1. Clearly, the value does not fall into the average category either.

Fuzzy logic will allow us to define membership functions that will let the hitter with 27 home runs be classified. For the sake of a future example, let us add one more hitter category, a hitter with less than 10 home runs will be considered a weak hitter. From the two previous definitions for good and average hitters and the new class of hitter, a fuzzy set membership function can be generated.

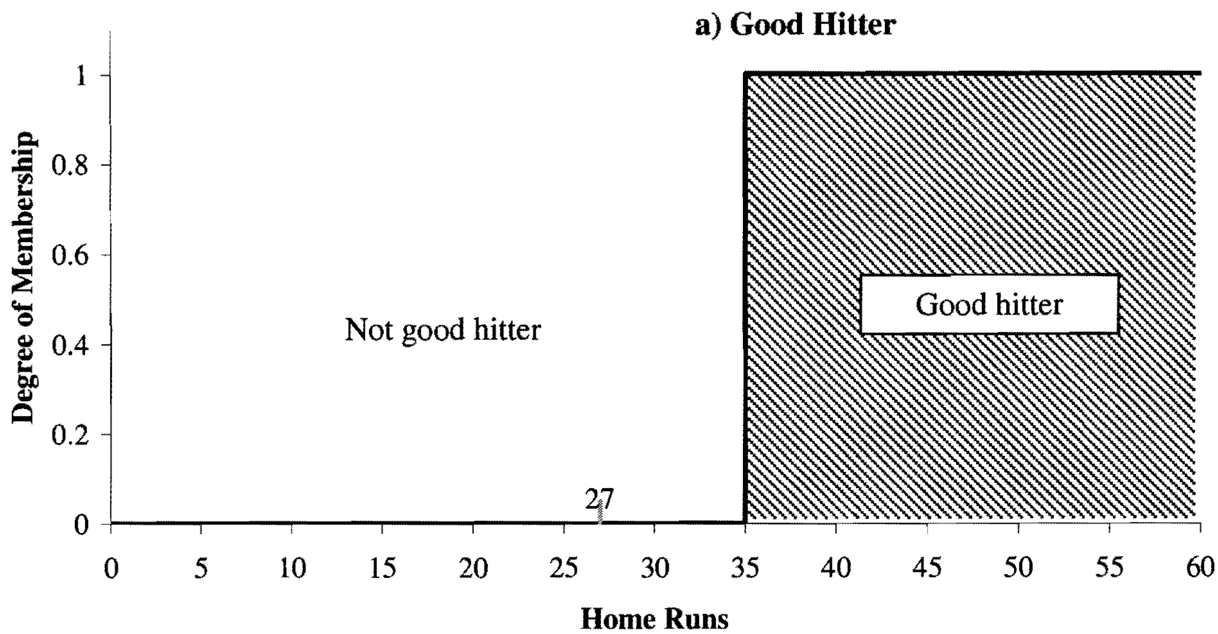


Figure 3.5 - Boolean Membership Function for Home Runs

By using a fuzzy set membership function, the degree of membership to each class can easily be defined by the three curves shown in Figure 3.6. As for the hitter in question, he will belong 53% to the average hitter class and 47% to the good hitter class. Observe how the 27 home runs are transferred to the y-axis to obtain the degree of membership where it intersects the membership functions for the average hitter and the good hitter. Obviously, he would not belong at all to the weak hitter class.

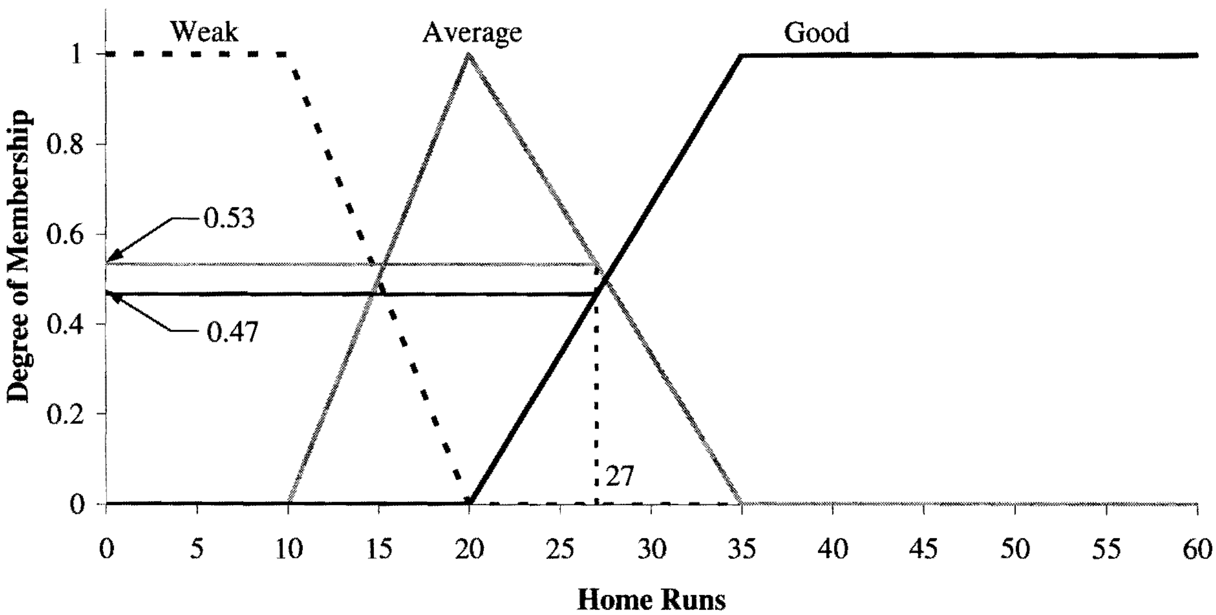


Figure 3.6 - Fuzzy Membership Function for Home Runs

There are four general shapes that define membership functions, Z, S, Λ, and Π. These shapes can be smooth curves or straight lines. The Z shape plateaus at one on the left, then slopes down to zero between two values and then is zero to the right of the higher value. The S shape is the reflection of the Z shape. The Λ shape is a triangle and the Π is a trapezoid. Figure 3.7 shows these generalized shapes. Any functions that create similar shapes can be substituted to create the desired membership function.

Additionally, these shapes can also be formed by sigmoid, Gaussian, and by double sigmoid curves. The shapes can also be asymmetric, that is having different ascending and descending slopes. The shape of the membership functions will be defined by "expert" input based on experience or empirical data.

Now that the use of fuzzy sets and membership functions has been described, the question arises of what to do with the information from the membership functions. Recall the example of the baseball player. The degree of belonging in each set has been determined: poor (0%), average (53%), and good (47%). Where the input (27 home runs) intersects a curve for a membership function, that input triggers a rule. That rule, in turn, will trigger an output. Now outputs must be developed. To start off simple, a good hitter will be given a rating of 3, an average player a rating of 2, and a poor player a rating of 1. The output can be expressed in graphical form as show in Figure 3.8.

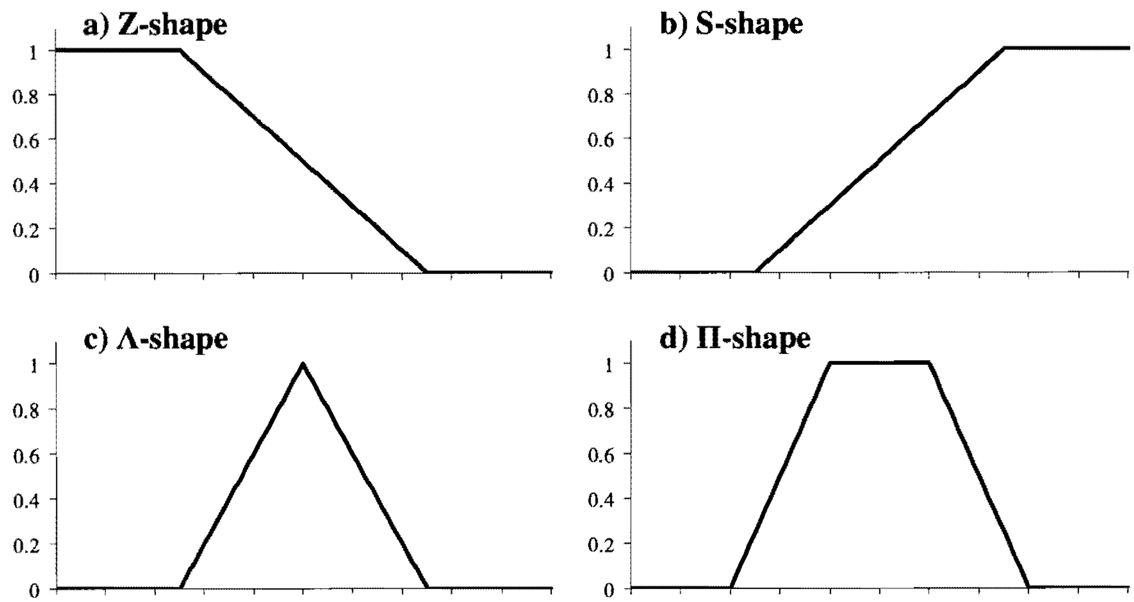


Figure 3.7 - Typical Shapes for Fuzzy Membership Functions

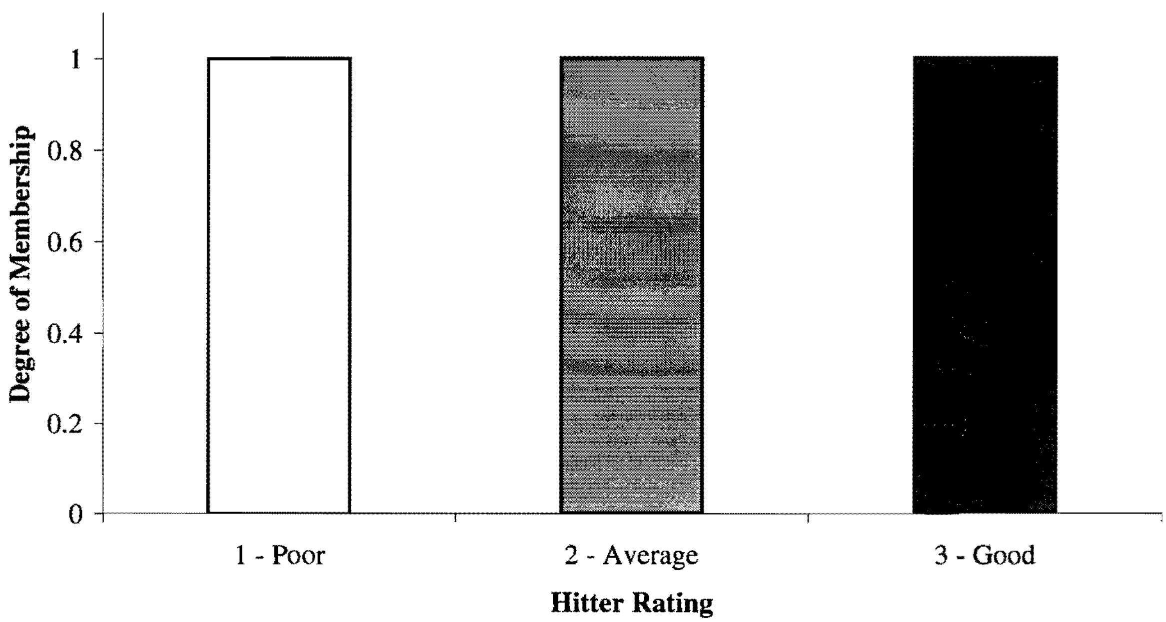


Figure 3.8 - Output for Baseball Example

Let us give each output possibility an equal weight of one. Now the degree of membership to each set can be applied to the output so the player can be rated by combining the outputs. Applying the 53% for the average rating and the 47% for the good rating, the output graph is transformed (Figure 3.9).

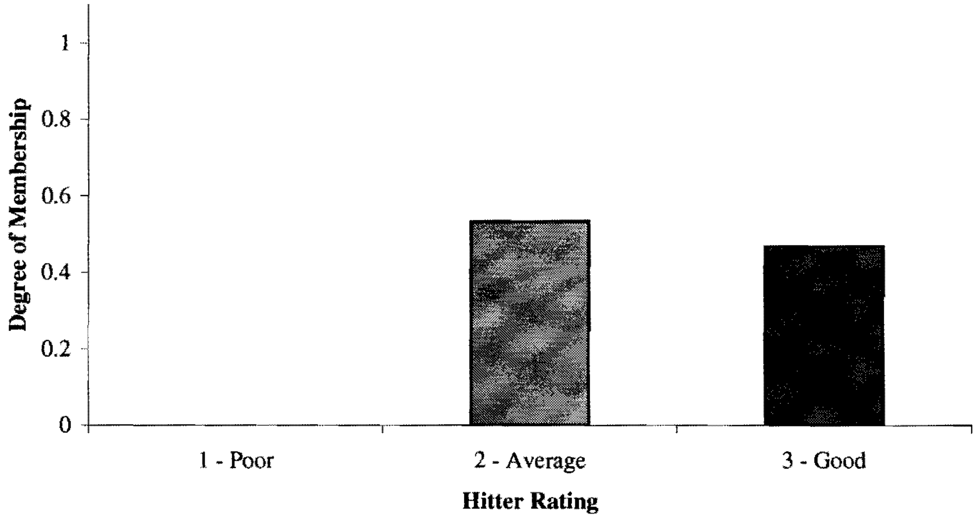


Figure 3.9 - Output Graph after Application of Inputs

Now the two values can be combined by taking the centroid of the areas. The result of the centroid is 2.47. This value was obtained by using the degree of membership as a weight and the output values as the numbers to be combined. Mathematically, the calculation is as following:

$$\frac{\sum w_i x_i}{\sum w_i} = \frac{(0.53 \times 2) + (0.47 \times 3)}{(0.53 + 0.47)} = 2.47 \tag{3.22}$$

This tells us that the player is between average and good, but slightly closer to being regarded as an average hitter than a good hitter. This process of simplifying the output is called defuzzification.

Realistically, the skill of a hitter will not be based on home runs alone. Now consider another input, batting average. Also, lets us add two more players such that the input matrix will look like Table 3.8.

Table 3.8 - Player Statistics Matrix

Player	Home Runs	Average
A	27	0.291
B	35	0.241
C	13	0.333

Now a second input membership chart must be developed. The concept of fuzzy operators will also be introduced at this point. Since there are now two inputs, a decision must be made to take the higher or lower value for the membership function that comes from the result of each input. Each rule will have two inputs (Home Runs and Average). Before, each rule had only one input (Home Runs). The membership functions for Average are shown in Figure 3.10.

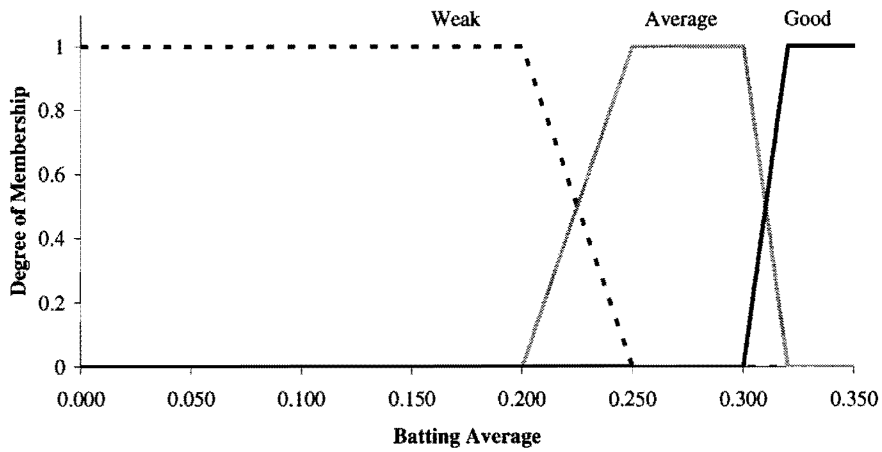


Figure 3.10 - Fuzzy Membership Function for Batting Average

Since there are multiple inputs, a way of deciding which input to use and which rules should be triggered become an issue. This is where a fuzzy operator will be introduced. When a decision must be made between two degrees of membership, the operators primarily used are *AND* and *OR*, or, as would be thought of in probability, intersection and union. Another basic operator used is the *NOT* (complement) operator. Two more variations of *AND* and *OR* are the product and algebraic sum. The product simply multiplies the degrees of membership and is classified as an *AND* method. The algebraic sum is P(a) + P(b) – P(a)*P(b) and is an *OR* method. In applying these operators, the *OR* function is equivalent to taking the maximum and the *AND* function is equivalent to taking the minimum of the numbers being compared (Mathworks, 2003).

Now the example of the baseball players can be further examined to show the fuzzy operators in use. Consider that each of the statistics shown in Table 3.8 will have an associated degree of membership. Each of those degrees of membership will either be in one or two classes as defined. Each degree of membership will then trigger one output. The rules have the following criteria:

1) If either Input 1 or Input 2, but not both, are equal to zero, the rule is triggered and only the non-zero input will be considered. The un-triggered input will have a null value and not be considered in any fuzzy operations.
2) If both Input 1 and Input 2 are zero, then the rule is not triggered.
3) If both Input 1 and Input 2 are non-zero then the rule is triggered and a decision is made which one to consider using a fuzzy operator.

Rule criteria are set by the user. The criteria set here are only for this example and will differ from case to case. Figure 3.11 shows how the two inputs for player A will trigger one or more of the three rules.

The first row corresponds, to rule one, the rule for a good hitter, rule two for average hitter and so forth. The first rule was only triggered by the 27 HR input. The second rule was triggered both by the 27 HR input and the 0.291 AVG. Rule three was never triggered. Graphs similar to those shown in Figure 3.11 can also be generated for players B and C.

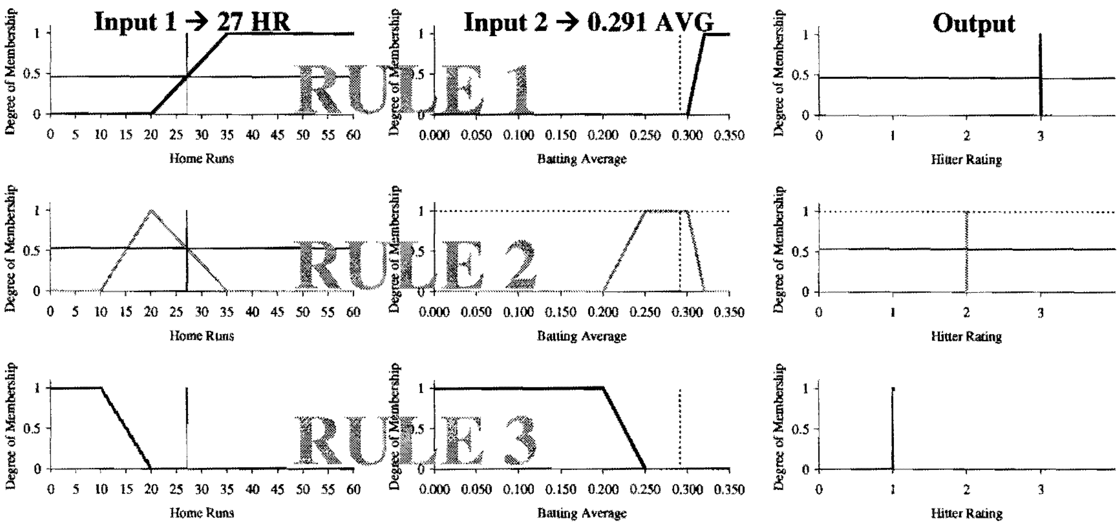


Figure 3.11 - Inputs and Outputs for Player A

Now a decision has to be made on which degree of membership value to use for the second output, the higher value or the lower value. This is where the fuzzy operators will be applied. For this example, the *AND* and *OR* operators will be used. If the *AND* operator is applied, the minimum value will be chosen, whereas using the *OR* operator, the maximum value will be selected. Table 3.9 shows the values for the triggered rules and the result of applying each operator for each of the three players.

When a rule is not triggered, the output is technically zero, but will not be considered when the fuzzy operators are applied (this was one of the rule criteria). Only non-zero values will have the *AND* or *OR* operators applied to them. The numbers in italics are the only difference in the example once the operators are applied. Now consider which of the two situations would qualify the player A as a better hitter, the *AND* case or the *OR* case. In this situation, the *AND* case would, since the result of the outputs is a centroid and the increase in weight for the average case would result in the centroid shifting down from the higher value of good that is supported by the 0.47 probability. Thus it may be more desirable to use the *AND* operator so that the player appears to be better. Conversely, the *OR* situation would imply that the player belongs more fully to the average category and in a sense would be a more fair estimate. For this example, both operators will be evaluated in the next step of combining the output. Additionally, the ratings for the hitters based solely on HR and AVG will be compared to the composite values.

The composite values will again be calculated using a centroid approach. This time there are four cases to compare, HR, AVG, *AND*(HR, AVG), and *OR*(HR, AVG). It was previously defined that a good player has a rating of three, an average player a rating of two, and a poor player a rating of one. Table 3.10 summarizes the player ratings obtained from the combining of the information. It is clear how the players will be ranked when rated solely on HR or AVG. However, the rating and ranking when the two statistics are looked at together is not so intuitive. Additionally, the ratings for each player change when the fuzzy operators are applied from what they were when simply based on one statistic. This rating process could be extended to include more players and/or more statistics.

Table 3.9 - Triggered Output Values and Fuzzy Operations

Player	Good	Average	Poor
Input 1 (Home Runs)			
A	0.47	0.53	--
B	1.0	--	--
C	--	0.30	0.70
Input 2 (Average)			
A	--	1.0	--
B	--	0.82	0.18
C	1.0	--	--
AND (minimum)			
A	0.47	*0.53*	--
B	1.0	0.82	0.18
C	1.0	0.30	0.70
OR (maximum)			
A	0.47	*1.00*	--
B	1.0	0.82	0.18
C	1.0	0.30	0.70

"--" denotes that the rule was not triggered

Table 3.10 - Comparison of Player Ratings and Rankings

Player	Rating	Ranking
Input 1 (Home Runs)		
A	2.47	2
B	3.00	1
C	1.30	3
Input 2 (Average)		
A	2.00	2
B	1.82	3
C	3.00	1
AND (minimum)		
A	2.47	1
B	2.41	2
C	2.15	3
OR (maximum)		
A	2.32	2
B	2.41	1
C	2.15	3

There are also other output curve options that are easy to use. Instead of having a concentrated mass for the outputs, the outputs could be represented by triangles, trapezoids, s- or z-curves, or by a function. In using such shapes and functions it would be necessary to develop appropriate widths for the output that would fairly weigh the results. One option is to have the same shape for each output, but with different centroids, and different heights where the heights are based on the inputs. In Figure 3.12 all triangles have the same base, but are truncated at different values. The bold black line contains the summed output areas. The final result will be the centroid of the area which is denoted by the dashed vertical line.

If certain rules were triggered more than once by different inputs, some of the areas would overlap, e.g. output three (O3) would have two trapezoids that fall under that triangle and thus more areas would be summed for that particular output.

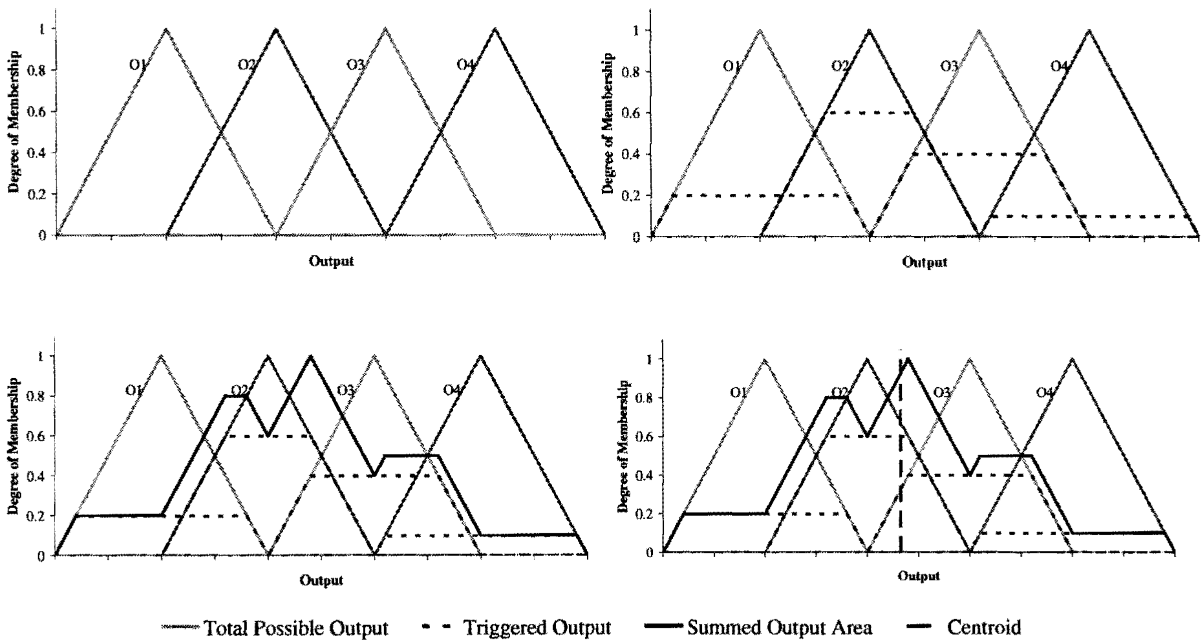


Figure 3.12 - One Alternative Form for Output Graphs

The rules used in the baseball example were quite simple in the sense that all three rules followed the same pattern. The three rules that were developed could have different conditions. For example if a player had more than 45 HR, qualifying him as a good hitter, only a severely low batting average would be applied to modifying his rating, say less than a 0.200 AVG. The charts shown in Figure 3.11 would change to accommodate this. Observe the change in the charts when this new criteria is applied to the previous data. An additional example would be to not penalize a hitter with an excellent average for not having as many home runs. Different membership functions could be used for input two based on the input for input one. There are any number of possibilities depending on what the user wants. Figure 3.13 shows the change that would occur to the previous rules once additional conditions are applied to the rules.

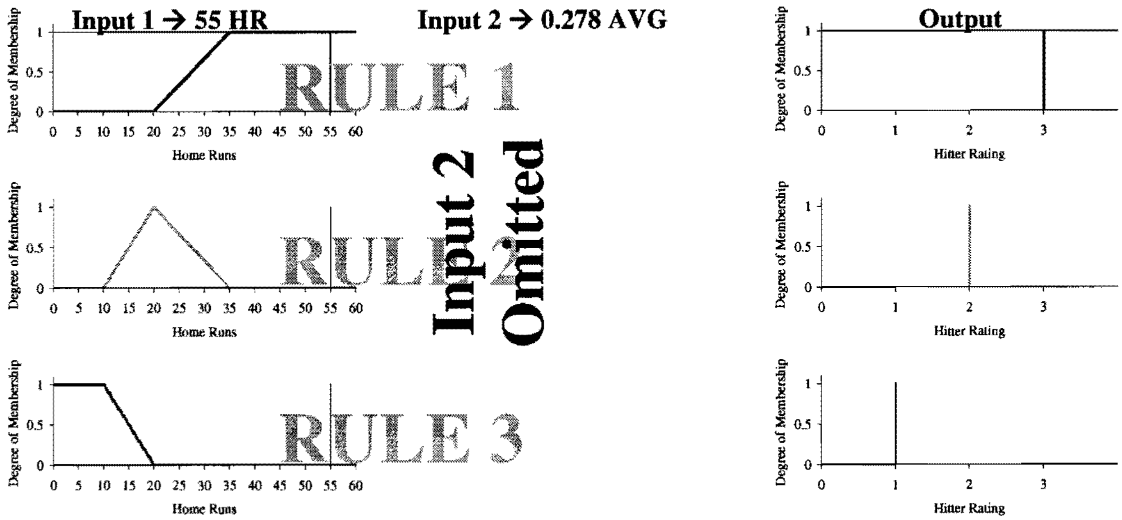


Figure 3.13 - Example of Conditional Rules

Fuzzy logic is a method that allows not only for the programming of rules to direct a certain outcome, but also for partial beliefs of membership of a given parameter to a set or multiple sets. By determining the degree of belonging to a set, outputs can be developed and then combined to obtain a composite view. The outputs obtained can represent anything from a final output to a weight that will be used in a latter step of data integration. The use of fuzzy logic is to set up a system of rules that will replace some of the guess work that a human being would have to perform in making a decision.

ADVANTAGES AND DISADVANTAGES OF METHODS

Statistical Weighted Average: The weighted average method is a simple way to fuse data from two or more sources. A representation of the uncertainty of the results is possible. This is important since a fused value cannot be perfect and must have some sort of uncertainty associated with it. However, this can also be viewed as a double-edged sword. The uncertainty will drop continually when more and more results are combined. This can lead to the assumption that the uncertainty is very low which may be true numerically, but not intuitively. Additionally, the use of this method requires that the results are in the forms of Gaussian distributions where means and standard deviations can be obtained for all data. Since it is fairly easy to calculate or assign reasonable standard deviations and means for most data, this method easily lends itself to data fusion. A method of using a statistical weighted average to fused NDT data from pavements will be introduced in the next chapters.

Bayesian Inference: The Bayesian inference method can be used to evaluate the most likely option from several choices. Additionally, it can easily be extended to generate weights that can then be used to fuse multiple options. The weakness of this method is that the generation of the conditional probabilities and prior probabilities can be tedious. The Bayesian method can also be used as a preprocessor to other methods of data fusion in which additional modification factors are developed that will aid in the weighting of data from multiple sources. A method of applying Bayesian inference to NDT pavement data will be presented in the next chapters.

Dempster-Shafer: The Dempster-Shafer method is an excellent way to combine data with overlapping sets. Dempster's rule for combining different data sets really adds to the attractiveness of this method. Ignorance in this method is represented by the confidence interval. The confidence interval gives a minimum and maximum probability of an event, referred to as belief and plausibility, respectively. If the discrete data sets are considered, combining the data becomes extremely easy. A degree of uncertainty can be obtained through fitting a distribution to the results via a numeric procedure or through the use of a distribution-fitting software. Although powerful, the Dempster-Shafer evidential reasoning method is excluded from use in the fusion of the NDT data. The input can be manipulated in such a way that another method can be used instead with easier computations to arrive at the exact same result.

Fuzzy Logic: The fuzzy logic method is an excellent method for developing weights that can be used to combine data or for combining the data itself. Fuzzy logic can easily incorporate sets of rules and partial belief functions to facilitate the development of the composite view of the data. Fuzzy logic can also be used to modify data as a pre-processing method for other fusion methods. To represent uncertainty in fuzzy logic, a composite value for the uncertainty must be derived. The equation for the uncertainty is not definitely defined. That is, the composite uncertainty function depends on how the outputs from the fuzzy procedure are combined. The same principles of error propagation that were used to define the uncertainty in the weighted average method can be applied to any fuzzy procedure that involves uncertainty. As the example shown in this chapter did not deal with uncertainty (i.e. home runs and batting average are know exactly), uncertainty cannot be defined. However, in the case of nondestructive test data for pavements, uncertainty plays an important role. The development of the uncertainty equations used for such data will be discussed in the chapter where fuzzy logic is applied to NDT pavement data. Perhaps the most important feature of the fuzzy logic is the way that it allows for the use of rules. These rules can easily be defined by the user and can be defined to consider any number of combinations of the input or inputs. The rules can be written to consider extreme or rare cases where special considerations must be taken. This allows the decision process to be standardized and not be subject to human interpretation which is likely to be inconsistent. The fuzzy logic method can also be applied in steps so that several levels of data combining can be performed.

ADDITIONAL LITERATURE ON DATA FUSION METHODS AND APPLICATIONS

In addition to the methods that were discussed in detail. Other data fusion methods exist that were not seen as applicable to this project. A brief description these methods will be included at this point.

Luo and Kay (1989) presented a summary of many fusion methods currently used in integrating intelligent systems. They referred to the different levels of fusion, such as fusing redundant information at a lower level and fusing complementary information at a higher level. They suggested which general methods should be used for which kind of fusion. The methods they covered are a weighted average, Kalman filter, Bayesian estimation using consensus sensors, multi-Bayesian, statistical decision theory, evidential reasoning, fuzzy logic and production rules.

Thomopoulos (1989) discussed how to combine data in a coherent manner. He detailed the architecture of data fusion by discussing the different levels of data fusion. Signal level, evidential level, and dynamic level fusion are discussed, as well as centralized and decentralized fusion. Signal level fusion can incorporate heuristic rules, correlation, or trainable networks (e.g. artificial neural networks). Evidential level fusion uses statistical models to describe the phenomenon being evaluated. The models can either be traditional or fuzzy. The outcome of this process is a local inference at the sensor level. Dynamic level fusion assumes that a mathematical model exists that describes the process for data collection from multiple sensors. Furthermore, it is also assumed that the data has a known transformation. A centralized approach would process the observations as a whole while a decentralized approach would process each sensor individually and then merge the processed data in the fusion center.

Other applications of data fusion include Krzysztofowicz and Long (1990), who formulated a Bayesian detection model for a distributed system of sensors. Additionally, Park and Lee (1993) used a fuzzy rule-based method for diagnosing and correcting signal faults as well as a method to check for sensor failures. Starr et al. (2000) discuss data fusion architectures, the blending of quantitative and qualitative methods, and applying data fusion for condition based maintenance. Osegueda et al. (2000) used several data fusion techniques to assess the damage of aluminum beams through fusion of modal strain energy differences. Methods that were used include averaging, likelihood ratios, Bayes statistics, and evidential reasoning.

CHAPTER FOUR

STATISTICAL WEIGHTED AVERAGE METHOD

INTRODUCTION

In Chapter 3 the concept and principles behind the weighted average method were presented. The implementation of the weighted average method in combining nondestructive pavement testing data will be presented in this chapter.

To apply the weighted average method, a mean and a standard deviation or a mean and a coefficient of variation are needed (Taylor, 1997). To represent this uncertainty it is desirable to average all the data from a particular device and consider that as the distribution for that device. In this study, the main concern is the fusion of the results, not the fusion of the raw data. Thus, when the term "device" is used, it will primarily refer to the entire process for using that device, including the reduction method. For composite reduction methods such as the joint inversion method where SPA and FWD data are reduced together, device will refer to the joint inversion method, or more simply as JIM (Abdallah *et al.*, 2003).

ADAPTATION OF NORMAL DISTRIBUTIONS FOR FUSION

The distributions developed represent the entire site and not the individual points. An adaptation of each site distribution must be made so that it can be used with the individual points. Suppose an FWD result, with a coefficient of variation of 12%, has a normal distribution of $\mathcal{N}(198, 24)$ and is derived from a set of twenty numbers. The first number in that list is 163. To adapt this distribution, the coefficient of variation is held constant, not the standard deviation. So the new distribution for the first test point for that sensor would be $\mathcal{N}(163, 12\% \times 163) = \mathcal{N}(163, 20)$. By using the coefficient of variation instead of the standard deviation, each individual value for a series of readings from a single device will have the same ratio between the inferred standard deviation for that point and the value for that particular test point. Figure 4.1 shows how the site data is transformed for one point. This same transformation can quickly and easily be done for any number of data points for a given site.

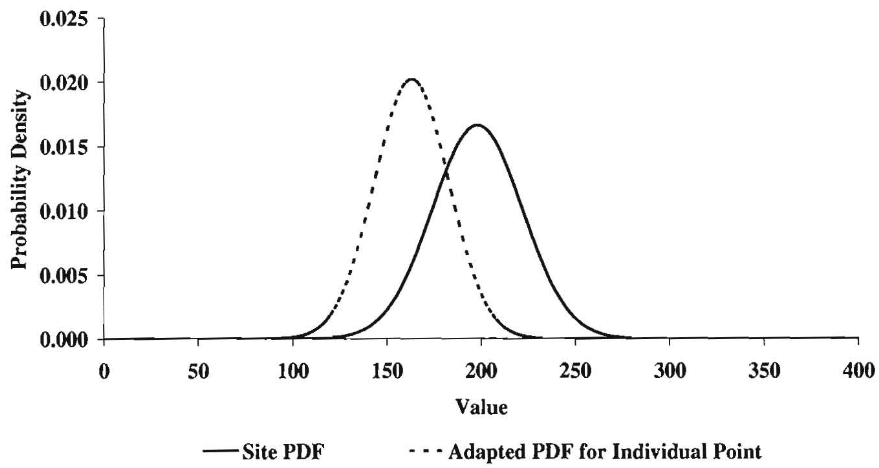


Figure 4.1 - Transformation of Distribution for Statistical Weighted Average

FUSION ARCHITECTURE

The site statistics will now be used to develop separate distributions for each of the individual test points. This procedure is fairly straightforward, but somewhat tricky. This tricky part is how to apply this method to a collection of data from a site where multiple devices were used to test it. Entertain the thought that there are many possibilities on how data from a site can be viewed. There are three classes where the weighted average method (and other methods in general) can be applied: point fusion, device fusion, and site fusion.

The concept of point fusion is quite simple, results from multiple devices must be aligned based on where they were tested; that is, the values that are to be fused must come from the different devices but from the same physical location. Figure 4.2 shows where point fusion would be possible for a set of data from a site tested with three devices. Keep in mind that in this discourse, device and method of reduction are often used synonymously.

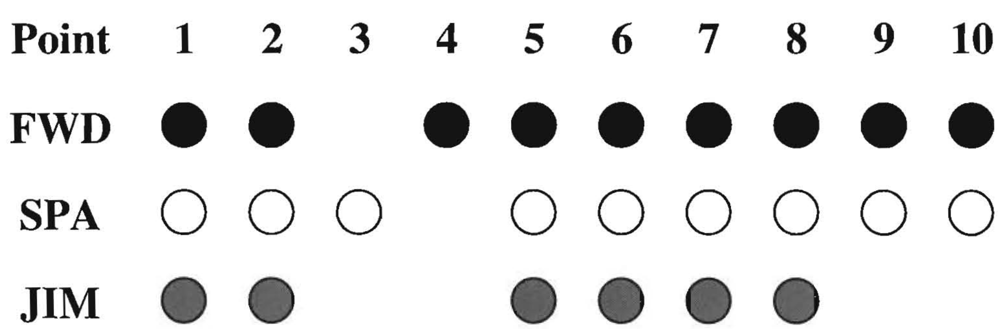


Figure 4.2 - Example of Available Data for Fusion

From Figure 4.2, point fusion will not be possible at every point. In addition, at points where it is possible, they may not be possible to the same degree. At some points results are not available because of, for example, a device malfunction, bad raw data that cannot be interpreted, or inconsistent or extremely outlying data. Since JIM data cannot exist without adequate FWD and SPA data, wherever one of those values is not available, the value from JIM is most likely unreliable. From Figure 4.2, all three devices can be fused at Points 1, 2, 5, 6, 7, and 8. Two devices can be fused at Points 9 and 10. Finally, at Points 3 and 4 no fusion is possible, as data was only available from one device. In that case, the value that will represent the properties at that point only comes from one device. Additionally, the points can be fused just using a simple average. This will provide some idea about how the data from the different devices can create a range of possibilities, but it will give a lesser weight to data that is more questionable (i.e., have a higher standard deviation) less. So each method of combining the data at a given point may have its advantages and disadvantages.

Device fusion and site fusion are fairly similar as both will provide some representation for the entire site. Device fusion will represent the site based on one device, while site fusion will represent the site based on multiple devices. Device fusion will be described first since site fusion is more complicated and often depends on the results of point fusion or device fusion. Referring to Figure 4.2 again, if device fusion were applied to the available data, three values would be obtained. Each device would provide a single value for the site. The limitation of device fusion is that an adequate representation of the site variation cannot be carried out. The site variability may be reduced during the fusion process to the point where the inherent variation is eliminated.

Site fusion can be performed in several ways. With the site fusion, the results of the other two forms of fusion have to be processed to obtain a final value for the site. One possibility is to fuse the data from point fused results. Other options are to fuse the device-fused data or fuse the simple device averages. Alternatively, all the points from all the sensors can be fused at once. Fusing point-fused results, fusing device-fused results, and fusing all will result in nearly exactly the same result. However, the method that makes the most sense is to take a simple average of the point fused data. This allows for more dependable readings at a given point to be weighted more while at the same time the variability within the site can be somewhat quantified. Figure 4.3 shows several alternatives for fusing the data from a particular site. These alternatives are more thoroughly explained where site fusion is specifically covered for real data.

DEVELOPMENT OF STATISTICAL WEIGHTED AVERAGE FUSION METHOD

Since the different possibilities of applying the weighted average method have been defined, an illustration of an actual application of this method to real data is in order. The data used in this example is from the parking lot at the TxDOT district office in El Paso. Three methods of data reduction were used in this case. One was using FWD, another was using SASW, and the third was using JIM. The coefficient of variation and average will be determined base on each device. The first step was to determine these coefficients, particularly the coefficients of variation for each parameter for all devices.

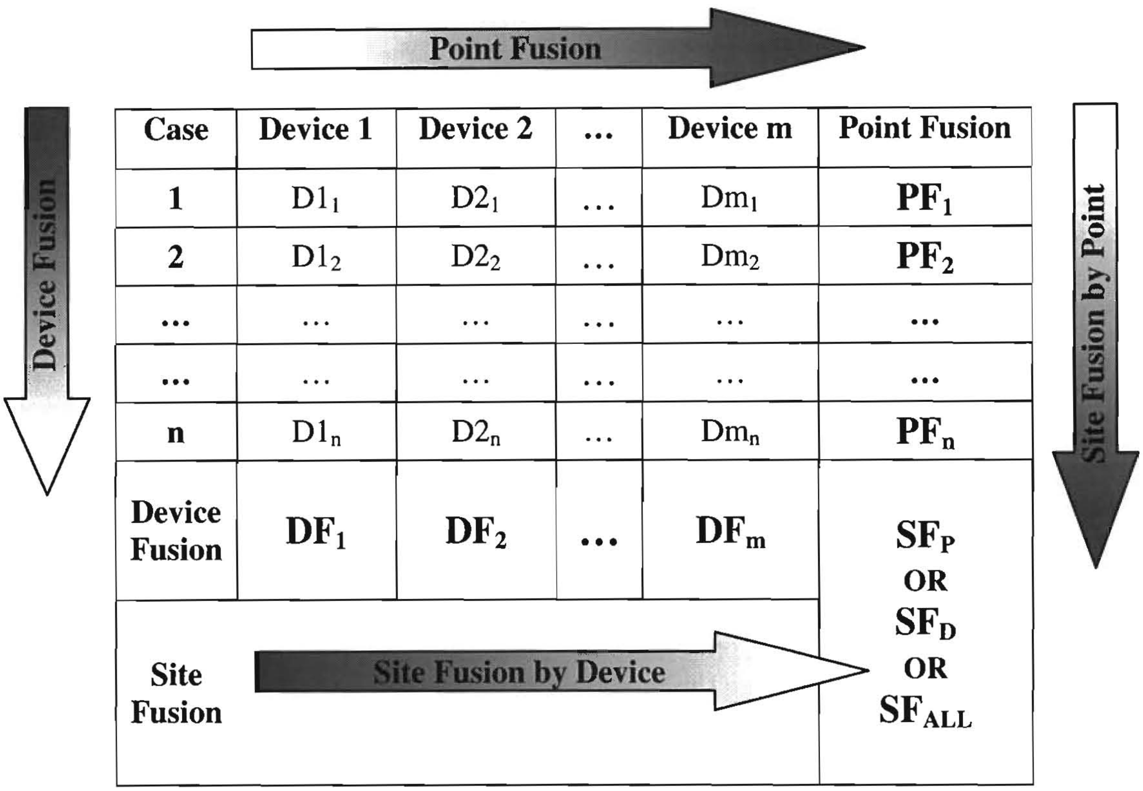


Figure 4.3 - Options for Fusion of Data

The second step is to find the weights by applying the coefficients of variation to transform the distribution for each point. The input data is included in Table 4.1. Equations 3.6 through 3.8 will be applied to generate the composite results.

Note that the standard deviation for the AC layer from Modulus is 0%. This is because the modulus values for the AC had to be fixed to properly reduce the data. This creates a false sense of uniformity of the modulus values for the site. In actuality the value that is fixed for the AC layer modulus is little more than a best guess by the user, and should not be weighted very highly. Fixing this modulus to another value would not have significantly affected the estimated moduli of the base and subgrade layers since the AC layer is thin. Therefore, in analyzing the data for this site an extremely high coefficient of variation (e.g. 1000000%) is considered for the AC layer. The value essentially makes the weight applied to this parameter so small that it will have virtually no impact when it is combined with the data from the other devices (methods). If for some reason the guess made by the user should be considered in the fusion of the data, an appropriate value can be assigned to the coefficient of variation to allow for the "guess" to have some influence on the fused result.

Table 4.1 - Data for Weighted Average Fusion

Test Pt.	AC Modulus (ksi)			Base Modulus (ksi)			SG Modulus (ksi)		
	Modulus*	SMART	JIM	Modulus	SMART	JIM	Modulus	SMART	JIM
1	500	647	672	68	102	60	21	20	12
2	500	632	636	43	62	68	22	20	12
3	500	636	652	54	139	92	19	17	11
4	500	556	501	55	45	73	17	21	10
5	500	496	547	73	107	72	18	20	10
6	500	659	605	119	182	105	19	21	10
7	500	732	681	152	169	93	19	21	11
8	500	508	549	156	242	91	19	15	11
9	500	689	734	189	180	92	17	13	10
10	500	590	619	89	132	80	16	17	9
11	500	404	412	43	81	64	14	16	8
12	500	621	584	78	146	99	16	21	9
13	500	571	527	112	117	82	18	20	10
14	500	455	490	77	167	92	17	14	10
15	500	609	601	110	127	104	17	14	10
16	500	521	486	62	126	92	17	21	9
17	500	563	527	78	44	72	17	21	9
18	500	552	526	59	57	69	17	17	9
19	500	543	603	47	103	64	16	17	9
20	500	668	655	108	113	102	18	14	10
Avg	500	583	580	89	122	83	18	18	10
StDev	0	82	80	41	51	15	2	3	1
CoV	10^6%	14%	14%	46%	42%	18%	10%	16%	11%

* Modulus value fixed for data reduction

For some of the base and subgrade measurements the values are conflicting, that is they are so far apart that the difference between the values is greater than either of the standard deviations for the values at that test point. Taylor (1997) mentions that this may be a problem with data and that the validity of the values must be double checked. Often times, for the conflicting points, one of the devices is already far from its own mean, hence, the value at that point for that particular device is already deviant from the mean for that device. Considering that the different methods take into account different inputs, some more tolerance can be allowed. But some values are quite discordant and therefore some additional reasoning must be used. For now, the incongruence of the numbers will be overlooked. A solution to this problem will be presented in a later chapter in which a hybrid method refines that data such that this problem is greatly reduced or even eliminated.

Next a table of weights can be generated that is a function of the coefficient of variation for the device and the modulus at a given point. Table 4.2 shows the weights for the site data was generated using Equation 3.6.

Table 4.2 - Weights for Fusion Process

Test Pt.	AC Weight (10^{-4})			Base Weight (10^{-3})			SG Weight (10^{0})		
	Modulus*	SMART	JIM	Modulus	SMART	JIM	Modulus	SMART	JIM
1	0.00	1.21	1.17	1.02	0.55	9.04	0.20	0.10	0.62
2	0.00	1.27	1.30	2.58	1.50	7.04	0.19	0.10	0.62
3	0.00	1.25	1.24	1.63	0.30	3.84	0.26	0.13	0.74
4	0.00	1.64	2.10	1.55	2.84	6.11	0.30	0.09	0.90
5	0.00	2.06	1.76	0.88	0.50	6.28	0.27	0.10	0.90
6	0.00	1.17	1.44	0.33	0.17	2.95	0.26	0.09	0.90
7	0.00	0.95	1.14	0.21	0.20	3.76	0.25	0.09	0.74
8	0.00	1.96	1.75	0.20	0.10	3.93	0.26	0.17	0.74
9	0.00	1.07	0.98	0.13	0.18	3.84	0.30	0.23	0.90
10	0.00	1.46	1.38	0.59	0.33	5.08	0.37	0.13	1.11
11	0.00	3.11	3.10	2.54	0.88	7.94	0.44	0.15	1.40
12	0.00	1.31	1.55	0.77	0.27	3.32	0.35	0.09	1.11
13	0.00	1.55	1.90	0.38	0.42	4.84	0.28	0.10	0.90
14	0.00	2.45	2.19	0.81	0.21	3.84	0.30	0.20	0.90
15	0.00	1.37	1.46	0.39	0.36	3.01	0.32	0.20	0.90
16	0.00	1.87	2.23	1.22	0.36	3.84	0.32	0.09	1.11
17	0.00	1.60	1.90	0.79	2.97	6.28	0.33	0.09	1.11
18	0.00	1.66	1.90	1.34	1.77	6.83	0.33	0.13	1.11
19	0.00	1.72	1.45	2.10	0.54	7.94	0.37	0.13	1.11
20	0.00	1.14	1.23	0.41	0.45	3.13	0.28	0.20	0.90

* CoV of 1000000% assigned

Now that the weights have been generated, the results can be fused. Table 4.3 shows the result of the point fusion. These results do not take into consideration the additional steps that can potentially be used to account for conflicting data.

Table 4.3 - Point Fusion Results

Test Pt.	Fused AC		Fused Base		Fused SG	
	Avg	CoV	Avg	CoV	Avg	CoV
1	659	10%	63	15%	15	7%
2	634	10%	61	15%	15	7%
3	644	10%	84	16%	13	7%
4	525	10%	63	16%	12	7%
5	520	10%	74	15%	13	7%
6	629	10%	110	15%	13	7%
7	704	10%	100	16%	14	7%
8	527	10%	97	16%	13	7%
9	711	10%	99	16%	12	7%
10	604	10%	84	15%	11	7%
11	408	10%	61	15%	10	7%
12	601	10%	98	15%	11	7%
13	547	10%	87	15%	13	7%
14	472	10%	93	15%	12	7%
15	605	10%	107	15%	12	7%
16	502	10%	88	16%	11	7%
17	543	10%	64	16%	11	7%
18	538	10%	66	15%	11	7%
19	570	10%	63	15%	11	7%
20	661	10%	104	15%	12	7%

The next type of fusion that can be performed on this data is device fusion. This will consist of fusing all the points for a particular parameter for a specific device only. Instead of averaging the results from one device for a particular parameter, the data will instead be combined using the weighted average fusion method. Since the weights were defined by the standard methods for calculating statistics, the same weights will be used as in the previous example where the point fusion was performed. This method will only yield nine results, one for each parameter for each device (three parameters and three devices). Table 4.4 shows the results for applying device fusion to the TxDOT parking lot data.

If the results in Table 4.4 are compared to those reflected in the last three rows in Table 4.1, the variability in the results has been greatly decreased. The only parameter that does not appear to be improved is the AC modulus using the MODULUS. The resulting coefficient of variation is high, but still considerably lower that the assigned 1000000%. It would make better sense to completely ignore this result anyway, since the value was just guessed.

Table 4.4 - Device Fusion Results

	AC Modulus (*ksi*)			Base Modulus (*ksi*)			SG Modulus (*ksi*)		
	Modulus*	**SMART**	**JIM**	**Modulus**	**SMART**	**JIM**	**Modulus**	**SMART**	**JIM**
Avg	500	558	558	64	78	78	17	17	10
StDev	1118034	18	17	7	8	3	0	1	0
CoV	223607%	3%	3%	11%	11%	4%	2%	4%	2%

* Modulus value fixed for data reduction and initial CoV set to 1000000%

Several methods for site fusion will be presented. Any number of combinations of point and device data can be transformed into site results. Table 4.5 shows several ways that the site fusion is applied to the TxDOT data.

Table 4.5 - Site Fusion Results

Site Fusion Method	AC		BS		SG	
	Avg	**StDev**	**Avg**	**StDev**	**Avg**	**StDev**
Average of Point Averages	554	52	98	32	15	1
Average of Device Averages	554	47	98	21	15	5
Average of Point Fusion	580	79	83	17	12	1
Average of Device Fusion	539	34	73	8	15	4
Average of All Points	554	76	98	42	15	4
Fusion of Point Averages	519	8	86	3	15	1
Fusion of Device Averages	581	57	86	13	12	1
Fusion of Point Fusion	558	12	76	3	12	0
Fusion of Device Fusion	558	12	76	3	12	0
Fusion of All Points	558	12	76	3	12	0

The ten methods for site fusion that are shown demand some special attention. Note how all the methods that solely rely on fusion, fusion of point fusion, fusion of device fusion, and fusion of all points, have the exact same results. If data from certain devices are missing (refer to Figure 4.2), these numbers will not be exactly the same, but fairly close. The fusion of fusions or fusion of all points greatly reduces the uncertainty, but in the process it eliminates the information that reveals the site variability. Other methods of site fusion are more practical for this purpose.

One more method that is subject to some scrutiny is the fusion of averages method, especially the fusion of point averages. Again, as a reminder, site fusion will give one value for an entire site. This is important to keep in mind because the reduction of uncertainty may not be the most important concern. First consider the fusion of point averages, specifically the fused subgrade average and coefficient of variation. It may initially appear that the near zero value for the standard deviation is indicating that the value is highly accurate and desirable. The reason for this value being so "accurate" must be examined to show a weakness for using a fusion of averages.

The fusion of the device averages seems to be less of a problem, since, the uncertainty of the individual devices is used in obtaining the final result. This method would be used where no data for a site was aligned (refer to Figure 4.2), such that the only way to obtain a fused result would be from fusion of fusions or fusion of averages. However, the drawbacks of fusion of fusion methods have already been discussed, so it is more advisable to use fusion of device averages than fusion of device fusion.

The methods that use averages of averages are not very desirable, because no consideration is placed on the variability that is inherent in the data and that some data may be more reliable than other data. There is no weighting of the data as all of the numbers put together using a regular average have an equal weight. Additionally, the averages come out the same (as in the fusion of fusions), but the standard deviations are different.

The method of averaging fused data is the most appropriate. The average of fused points allows for a representation of the site variability while at the same time fusing the point data based on weights that will put more credence into more reliable values. The average of fused devices would be advisable when point fusion is not feasible due to data not being aligned. Instead of having a value from each device, overall site properties can be obtained. Naturally, if the values are very discordant, some examination of the input should be made to verify that no values that are overly incongruent are being combined. Discrepant results can be dealt with appropriately by using the hybrid method that will be discussed in Chapter 6 once the appropriate foundation has been laid.

Other considerations are to use a device reliability value instead of site statistics, if the site statistics drops below a certain allowable minimum for repeatability. That is, say that a device is known to be 90% reliable and the site statistics have a coefficient of variation of 20%, thus the 20% value will be acceptable to use. However, what if the site statistics result in a 3% coefficient of variation? This will mean that a confidence is placed on the results by the site statistics that is not supported by the device capabilities. In other words, the device is said to be more reliable than it actually is. At this juncture it would be advisable to use a coefficient of 10% (100% – 90%) instead of 3% in fusing the data.

THE WEIGHTED AVERAGE ALGORITHM

The background and mathematics for the weighted average method have been presented in the previous section of this chapter. As this method is fairly straightforward, the only difficulty that may arise is which numeric value to use for the coefficient of variation to establish the weights. There are three basic options to establish the coefficient of variation. The first option is to use the site coefficient of variation for a particular sensor. The second is to use a user defined coefficient of variation. This user defined coefficient of variation can be defined by a variety of factors such as device repeatability, prior information regarding the site data, or trust in the measurement due to operator skill or competency. The third option would be to omit the data for a given sensor. This could be done simply by the user defining a very high coefficient of variation.

The procedure is very simple for using this algorithm. The only preprocessing of the data that is necessary is just to obtain the modulus values from their respective methods and then align the data so that only data from the same test point will be combined. After that, the desired options for the coefficient of variation must be briefly evaluated by the user to decide which method to use. For most purposes, it is acceptable and desirable to simply use the site coefficients of variation from each device. Figure 4.4 shows the procedure used for the weighted average method. The site fusion options that are shown on the flowchart are the same ones that were listed in Table 4.5. Furthermore, as was described earlier, the most desirable method to use is to fuse the point data and then to use averages for the sites from that point on. The output from using fusion for the points and an average for the site will be a list of values for the points and a mean and standard deviation for the site average.

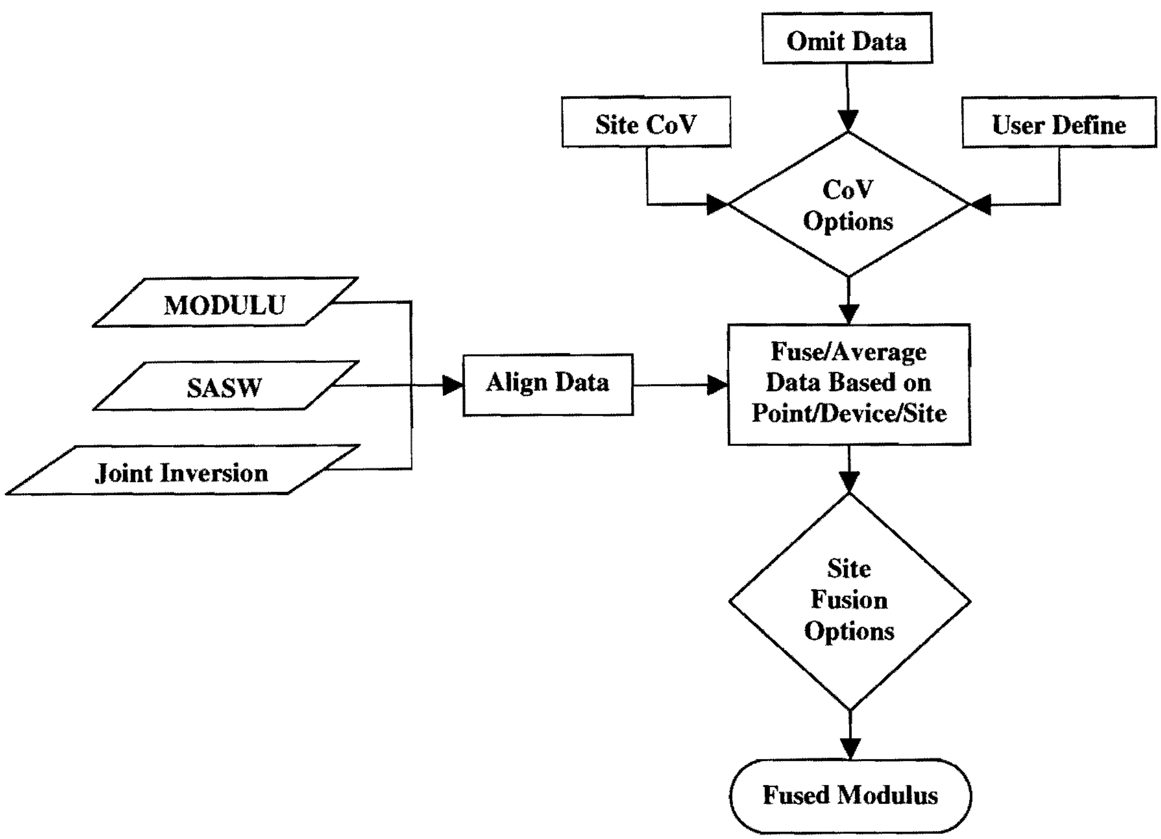


Figure 4.4 - Statistical Weighted Average Algorithm

The weighted average method is a simple straightforward method of fusing data from one or more sources. The over-reduction of uncertainty that arises when several values are combined may be a problem. As such, this method should be used in concert with a regular average such that a more realistic uncertainty of a given site can be found. Furthermore, if the numbers that are being combined are contradictory or too distant, it is advisable to verify the validity of the data or to flag any values that are formed from discordant data such that they may be identified.

This method can easily be incorporated into a hybrid method of data fusion. This is because the weighted average method is simple to apply once a representation of uncertainty is obtained in the form of a standard deviation (or coefficient of variation). Modifiers can easily be placed to reduce or increase this uncertainty based on a number of criteria, like deviation from the mean for example. A more detailed discussion on how this method can form a hybrid method with other methods, namely fuzzy logic, will be discussed in Chapter 6.

CHAPTER FIVE

FUZZY LOGIC METHOD

INTRODUCTION

The fuzzy logic method can be applied to nondestructive testing data to both combine and filter data. Through the use of fuzzy rules, numerous checks on the input can be made.

The examples presented in Chapter 3 used inputs (homeruns and batting average) to generate a rating for a baseball player. This was done by converting each input to an output that represented a weighting value that could be used to combine data from any of the rules that were triggered. A similar approach will be taken in applying this method to NDT data.

DEVELOPMENT OF FUZZY LOGIC METHOD

The most crucial problem experienced in the applying of fuzzy logic to NDT data for this project was defining the inputs. The problem came up on how to use, to obtain or to convert the moduli to weights that could then be used to combine data from two or more different devices. Another option would be to make discrete choices based on the moduli themselves or based on deviations from the mean of those values. In this case, no values are combined, but a choice is made between the given values.

The procedure that will be covered is a method of obtaining weights from the moduli themselves, either directly or from some preprocessing of the data. An additional input can be generated from the reliability of the device or the variability in the data. This was handled by adding an option for an additional modification factor.

This method directly uses the moduli and will incorporate a procedure similar to that of the baseball player examples. Appropriate criteria will be set to define very low, low, average, and high quality values for moduli and then weights can be determined from them. Very Low is defined as minus two standard deviations from the mean, low as minus one standard deviation, average at zero, and high at plus one standard deviation. The corresponding output will again be weights as in the baseball example from Chapter 3; however, there are three parameters that must be analyzed separately (i.e. AC, base, and subgrade moduli). In illustrating this method the data from the TxDOT parking lot shown in Table 4.1 will be used.

To apply this method, membership functions must be developed as the first step. The average value will be determined from an average of all points from all sensors for a given parameter. Similarly, the standard deviation of all the values for a particular parameter will be used. These parameters are calculated for asphalt concrete (AC), base, and subgrade moduli. Table 5.1 shows the average and standard deviation for the site when all values are included.

Table 5.1 - Site Averages Used to Develop Membership Functions

Modulus (ksi)	AC	Base	SG
Average	554	98	15
StDev	76	42	4
CoV	14%	43%	28%

To make the input more robust, each of the averages and standard deviations were rounded. The AC modulus was rounded to the nearest 25, the base to 10, and the subgrade to 5. These values were arbitrarily chosen for this example. Additionally, the boundaries for the membership functions were defined by standard deviations from the mean. The modified values are show in Table 5.2.

Table 5.2 - Rounded Site Data and Membership Boundaries

Parameter	Modulus (ksi)		
	AC	Base	SG
Average	550	100	15
StDev	75	40	5
Set	Membership Boundaries (*ksi*)		
Very Low	400	20	5
Low	475	60	10
Average	550	100	15
High	625	140	20

The shapes of the functions can be described in terms of the general forms introduced in Chapter 3. The "Very Low" set is Z-shape, the "High" set is S-shape, and the two middle sets, "Low" and "Average" are Λ-shapes (triangular). Figure 5.1 graphically shows these shapes for the three moduli.

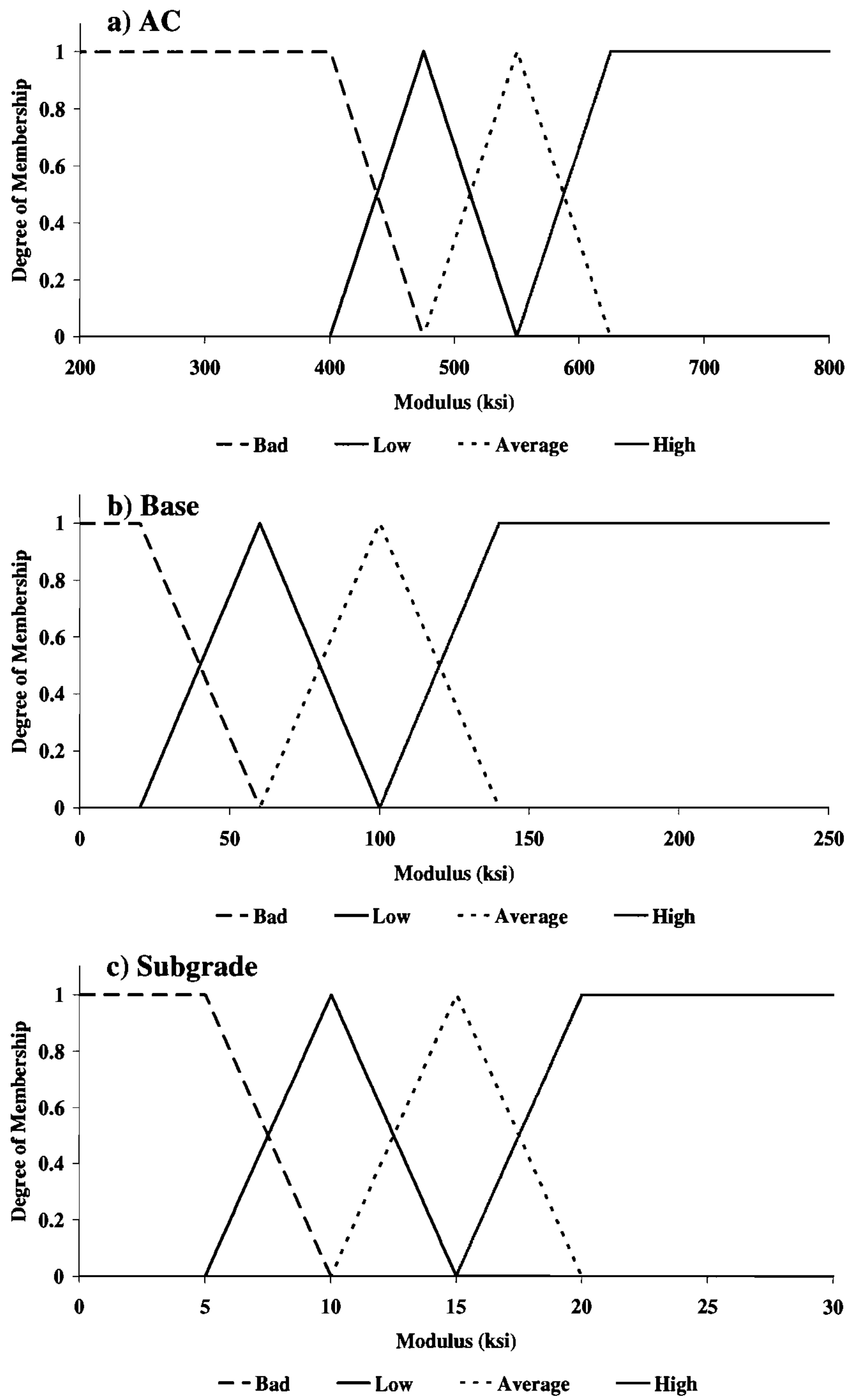


Figure 5.1 - Membership Functions for TxDOT Data

Each of the diagrams in Figure 5.1 represents four rules that can be triggered depending on the inputs. Each series (very low, low, average, high) will represent an individual rule, and each graph in Figure 5.1 could further be decomposed into the four categories of material quality.

Next the outputs for each rule must be developed. To quantify the weight for each triggered rule, values will be assigned based on the importance of the rule triggered. With the baseball example, the better the statistic corresponded to a higher weight and this was directly used to rate the player. In this case, the output will be used to generate a weight for the values from each sensor by using the centroid approach. Once the weights have been generated, it is just a matter of taking a weighted average to fuse the results.

The procedure for fusing the data has been described, but now the weights must be defined for each parameter. For layers that may be more crucial and a conservative estimate is more desirable, the weights for low and very low should be increased. Average values should have one of the highest weights, as they are in the expected range. The weight for high should be lower than the average as it is not really expected, but a determination of how high it should be relative to low and very low depends on how crucial the data from that layer is. Tables 5.3 through 5.5 show the results of inputting the data from Case 1 from Table 4.1 into the fuzzy procedure. The four rules are described in the table as $f_n(x)$ and a matrix of values is produced. The preliminary weights are also shown for all three layers. Note how each layer may be very low quality in one area, average in another, and excellent in the third.

Note how the weights for the poor quality of the base and subgrade are larger than the AC layer and the subgrade larger than the base. This was to allow for the fused base and especially subgrade modulus values to be much more conservative when there was input that supported a lower value. Since the AC values tend to be much closer, in general, among different methods, the skewing of the data toward the conservative side is not as necessary. The way the weights have been defined takes into consideration two items: 1) the deviation of individual point values from the site mean and 2) the greater caution placed on layers that depend on the degree of conservativeness that is desired.

Table 5.3 - AC Fuzzy Membership Values and Weights

Parameter	**Very Low**	**Low**	**Average**	**High**	**Sum**
	$f_1(x)$	**$f_2(x)$**	**$f_3(x)$**	**$f_4(x)$**	**Σ**
Modulus – 500 ksi	0	0.67	0.33	0	1
SMART – 647 ksi	0	0	0	1	1
JIM – 672 ksi	0	0	0	1	1
w_1	1	2	3	2.5	--
w_2	1	2	3	2.5	--
w_3	1	2	3	2.5	--
$w_1 \times f_n(x)$	0.0	1.3	1.0	0.0	2.3
$w_2 \times f_n(x)$	0.0	0.0	0.0	2.5	2.5
$w_3 \times f_n(x)$	0.0	0.0	0.0	2.5	2.5

Table 5.4 - Base Fuzzy Membership Values and Weights

Parameter	Very Low	Low	Average	High	Sum
	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	Σ
Mod 5.1 – 68 ksi	0	0.8	0.2	0	1
SMART – 102 ksi	0	0	0.95	0.05	1
JIM – 60 ksi	0	1	0	0	1
w_1	1.5	2	3	2.5	--
w_2	1.5	2	3	2.5	--
w_3	1.5	2	3	2.5	--
$w_1 \times f_n(x)$	0.0	1.6	0.6	0.0	2.2
$w_2 \times f_n(x)$	0.0	0.0	2.9	0.1	3.0
$w_3 \times f_n(x)$	0.0	2.0	0.0	0.0	2.0

Table 5.5 – Subgrade Fuzzy Membership Values and Weights

Parameter	Very Low	Low	Average	High	Sum
	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	Σ
Mod 5.1 – 21 ksi	0	0	0	1	1
SMART – 20 ksi	0	0	0	1	1
JIM – 12 ksi	0	0.6	0.4	0	1
w_1	1.5	2	2.5	2.25	--
w_2	1.5	2	2.5	2.25	--
w_3	1.5	2	2.5	2.25	--
$w_1 \times f_n(x)$	0.0	0.0	0.0	2.3	2.3
$w_2 \times f_n(x)$	0.0	0.0	0.0	2.3	2.3
$w_3 \times f_n(x)$	0.0	1.2	1.0	0.0	2.2

If all values (from each method) at a particular point are well below (or above) the mean, the weight that each value will be assigned will be relatively close to one another. Thus, points where all data support low (or high) values will not be classified as an outlying point. This is because the data for that point supports the low or high modulus, although the overall data from the site may not. If there are conflicting results at a point, it may be advisable to verify or remove a potential outlier. If this is not possible, the user should decide which value is most likely to be the outlier. One suggestion is to view the site data in making this determination.

Once the weights have been defined, modification factors can be applied to the process that will take into account special circumstances. For example, while reducing data from MODULUS, sometimes it is necessary to fix the modulus for the AC layer. As with the weighted average method, this again will be just a guess value and little or no credence may be placed into it. Thus a modification factor can be applied to either eliminate or weigh those values less. Additionally, device repeatability can be included as an additional modification factor, as can the program sensitivity. This will allow for a composite modification factor to be obtained by multiplying a series of corrective factors. These will modify the preliminary weights obtained from the triggering of the rules. Say that the modulus of the AC layer is 50% (0.50) credible by reason of construction data, past testing results, or other pertinent information. The weight would be reduced by multiplying the outputs of the fuzzy procedure by this number. Table 5.6 shows the modification factors for the TxDOT data. Only the AC modulus was affected by these factors. The base and subgrade weights were accepted with no additional modification. The reason that the AC modulus values required the modification factors, k_m, is because the AC modulus was fixed when the data was reduced using MODULUS.

Table 5.6 - Modification Factors for Fuzzy Logic

Method	Device Modification Factor		
	AC	Base	Subgrade
MODULUS	0	1	1
SMART	1	1	1
JIM	1	1	1

The modification factors, k, must now be appropriately applied to the parameter or parameters that require attention. Table 5.7 shows the applied modification factors to Table 5.3. Note how the weight for the MODULUS was eliminated in this case.

Table 5.7 - AC Fuzzy Membership Values and Weights with Modification Factors

Parameter	Very Low	Low	Average	High	Sum
	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	Σ
MODULUS – 500 ksi	0	0.67	0.33	0	1
SMART – 647 ksi	0	0	0	1	1
JIM – 672 ksi	0	0	0	1	1
w_1	1	2	3	2.5	--
w_2	1	2	3	2.5	--
w_3	1	2	3	2.5	--
$w_1 \times f_n(x) \times k_1$	0.0	0.0	0.0	0.0	0.0
$w_2 \times f_n(x) \times k_2$	0.0	0.0	0.0	2.5	2.5
$w_3 \times f_n(x) \times k_3$	0.0	0.0	0.0	2.5	2.5

The final composite weights and each parameter for each device are shown at the last three rows of the last columns of Tables 5.4, 5.5, and 5.7. Table 5.8 presents a summary of those values and of the input moduli that were used to obtain those weights. Additionally, the non-weighted averages are shown.

Table 5.8 - Fused Results for Case 1 Using Fuzzy Logic

Method	AC		Base		Subgrade	
	Modulus	Weight	Modulus	Weight	Modulus	Weight
MODULUS	500	0.0	68	2.2	21	2.3
SMART	647	2.5	102	3.0	20	2.3
JIM	672	2.5	60	2.0	12	2.2
Fused	660	--	80	--	18	--
Average	606	--	77	--	18	--

This method was applied to each of the twenty test points from the TxDOT parking lot. A comparison between the simple average values and the fused values is in order. Another comparison will be between the initial (unfused) values and the fused results so that the most influential value can be observed. Figure 5.2 shows the variation of the input and the two outputs for the TxDOT data. All fused values lie between the minimum and maximum values at a particular point. It appears that the average is not that different from the fused values, with the exception of the AC modulus. This is due to the fact that the values were not very deviant from the mean for either the base or subgrade. However, for the AC modulus the results from MODULUS were weighted out. Any points with outliers far from the mean would be weighted considerably less, especially if more membership classes were used (e.g. very low, very good, excellent). This would result in more of the points on the fringes of the modulus range to be looked at with more scrutiny due to their deviation from the mean. The numeric values that are plotted in Figure 5.2 are shown in Table 5.9 in the plot.

It was previously mentioned that the difference between the average and fused results was small. The ranges for the percent of the average differing from the fused value are shown in Table 5.10. The percent difference is higher for the AC modulus because the value from MODULUS was omitted from the calculation. The average was very close to the fused value in nearly all of the cases for the base and subgrade because the data was fairly consistent throughout the site and the deviation from the mean was similar enough that the weights that were outputted from the fuzzy process were about the same.

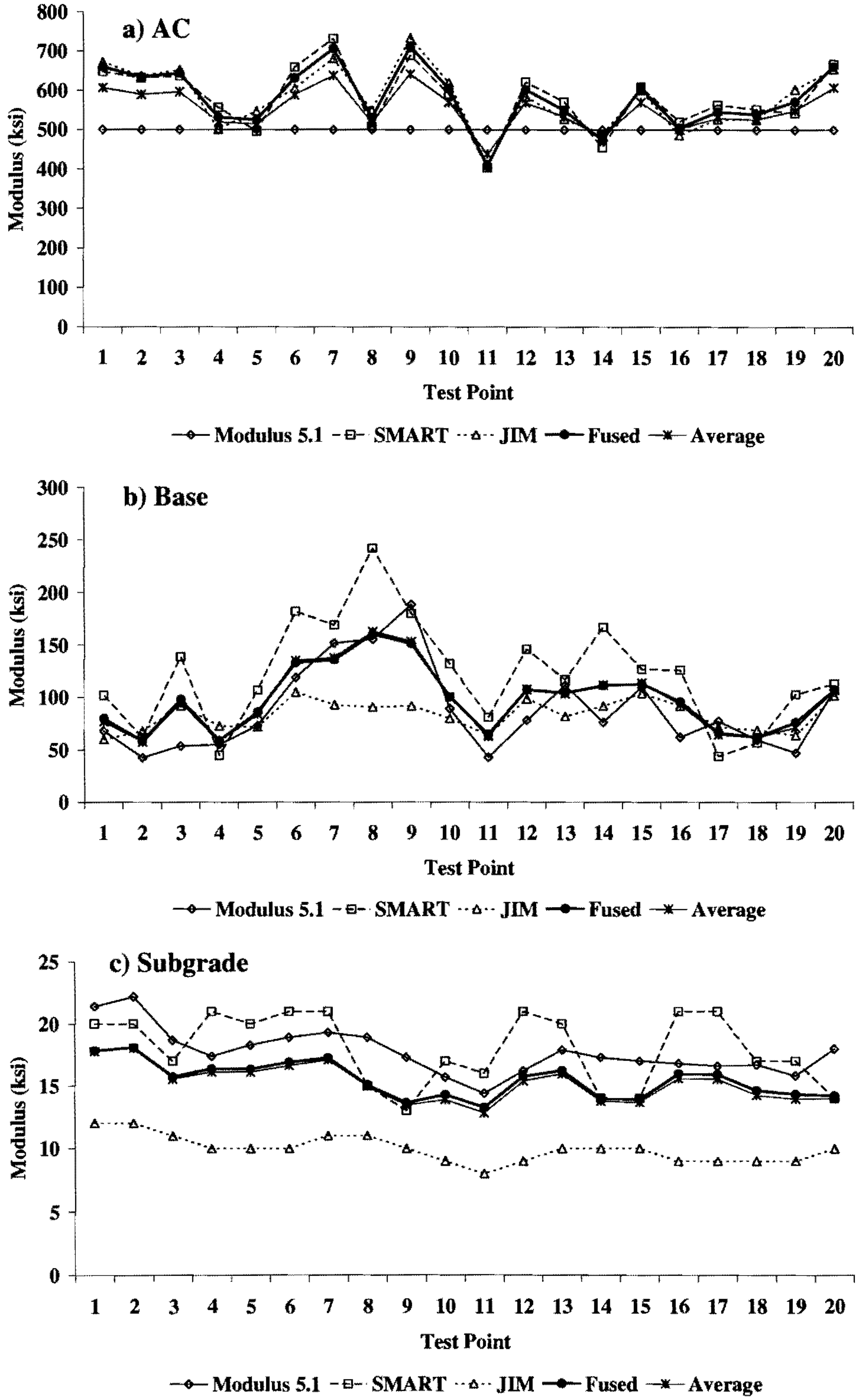


Figure 5.2 - Fusion Results for TxDOT Data Using Fuzzy Logic

Table 5.9 - Results of Point Fusion Using Fuzzy Logic

Test Pt.	M*	S*	J*	F*	M	S	J	F	M	S	J	F
	AC				BASE				SG			
1	500	647	672	660	68	102	60	80	21	20	12	18
2	500	632	636	634	43	62	68	59	22	20	12	18
3	500	636	652	644	54	139	92	98	19	17	11	16
4	500	556	501	532	55	45	73	59	17	21	10	16
5	500	496	547	525	73	107	72	86	18	20	10	16
6	500	659	605	631	119	182	105	133	19	21	10	17
7	500	732	681	707	152	169	93	136	19	21	11	17
8	500	508	549	531	156	242	91	160	19	15	11	15
9	500	689	734	712	189	180	92	151	17	13	10	14
10	500	590	619	604	89	132	80	101	16	17	9	14
11	500	404	412	408	43	81	64	65	14	16	8	13
12	500	621	584	602	78	146	99	107	16	21	9	16
13	500	571	527	550	112	117	82	104	18	20	10	16
14	500	455	490	475	77	167	92	111	17	14	10	14
15	500	609	601	605	110	127	104	113	17	14	10	14
16	500	521	486	505	62	126	92	96	17	21	9	16
17	500	563	527	546	78	44	72	66	17	21	9	16
18	500	552	526	540	59	57	69	62	17	17	9	15
19	500	543	603	572	47	103	64	76	16	17	9	14
20	500	668	655	662	108	113	102	107	18	14	10	14
Avg	500	583	580	582	89	122	83	99	18	18	10	15
StDev	0	82	80	78	41	51	15	30	2	3	1	1
CoV	10^6%	14%	14%	13%	46%	42%	18%	31%	10%	16%	11%	9%

* M = MODULUS, S = SMART, J = JIM, F = Fused

Table 5.10 - Fusion and Average Comparison

$\frac{\lvert x_{Fused} - x_{Average} \rvert}{x_{Fused}} \times 100\%$	Error Values		
	AC	Base	Subgrade
Min	0.6%	0.1%	0.2%
Max	9.9%	6.3%	3.5%
Avg	5.1%	1.8%	1.7%
StDev	2.9%	1.6%	0.9%
CoV	56.0%	86.4%	49.9%

REPRESENTATION OF UNCERTAINTY OF FUZZY LOGIC METHOD

In representing uncertainty with the weighted average method, the principles of error propagation were applied (Taylor, 1997). The same principles can be used with fuzzy logic approach to derive an equation to represent errors. However, due to the complexity of the fuzzy logic process, some simplifications and assumptions must be made. Firstly, the equations used must be defined symbolically. The equation for the final fused value is simply a weighted average where the weights come from the fuzzy sets. The fused value, μ_F, found using the fuzzy procedure can be defined symbolically by Equation 5.1

$$\mu_F = \frac{\sum k_{m_i} w_i x_i}{\sum k_{m_i} w_i} \tag{5.1}$$

where the k_{m_i}'s refer to the modification values and the w_i's and x_i's are simply the weights from the fuzzy sets and the values that are to be combined, respectively. Technically, the weights are a function of the input values (the x_i's). Since the number and shape of the membership functions that the weights are obtained from vary, the equations can rapidly become overly complex. The primary uncertain parameter is the modulus. Predicting the uncertainty for the fused numbers is even more complex as no definitive method to represent their uncertainty exists. Thus only the uncertainty in the values themselves will be considered and the uncertainties in the modification factors will not be considered. The same principles that were described in Chapter 3 for summations of variables are applied to Equation 5.1. The resulting uncertainty equation will be as follows

$$\sigma_F = \sqrt{\frac{\sum \left(k_{m_i} w_i \sigma_i\right)^2}{\left(\sum \left(k_{m_i} w_i\right)\right)^2}} \tag{5.2}$$

Table 5.11 shows the initial uncertainty obtained from the site statistics and the uncertainty of the fused values. The uncertainties are given in terms of standard deviation. All inputs do impact the uncertainty. As such, the composite uncertainty is not necessarily lower than its component uncertainties. The modification values as well as the values of the weights impact the resulting uncertainty as represented in Equation 5.2.

Table 5.11 - Uncertainty of Fused Results

Test Pt.	M*	S*	J*	F*	M	S	J	F	M	S	J	F
	AC				BASE				SG			
1	N/A	91	93	65	31	43	11	20	2.2	3.2	1.3	1.4
2	N/A	89	88	62	20	26	12	11	2.3	3.2	1.3	1.4
3	N/A	89	90	63	25	58	16	22	2.0	2.7	1.2	1.2
4	N/A	78	69	53	25	19	13	11	1.8	3.4	1.1	1.4
5	N/A	70	75	52	34	45	13	21	1.9	3.2	1.1	1.3
6	N/A	93	83	62	55	76	18	30	2.0	3.4	1.1	1.4
7	N/A	103	94	70	70	70	16	32	2.0	3.4	1.2	1.4
8	N/A	71	76	53	72	101	16	40	2.0	2.4	1.2	1.2
9	N/A	97	101	70	87	75	16	37	1.8	2.1	1.1	1.0
10	N/A	83	85	59	41	55	14	24	1.6	2.7	0.9	1.2
11	N/A	57	57	40	20	34	11	15	1.5	2.6	0.8	1.1
12	N/A	87	80	59	36	61	17	23	1.7	3.4	0.9	1.3
13	N/A	80	73	54	51	49	14	25	1.9	3.2	1.1	1.3
14	N/A	64	68	47	35	70	16	26	1.8	2.3	1.1	1.1
15	N/A	86	83	60	51	53	18	25	1.8	2.3	1.1	1.1
16	N/A	73	67	50	29	53	16	21	1.8	3.4	0.9	1.4
17	N/A	79	73	54	36	18	13	15	1.7	3.4	0.9	1.4
18	N/A	78	72	53	27	24	12	12	1.7	2.7	0.9	1.2
19	N/A	76	83	56	22	43	11	20	1.7	2.7	0.9	1.2
20	N/A	94	90	65	49	47	18	23	1.9	2.3	1.1	1.1

* M = MODULUS, S = SMART, J = JIM, F = Fused

THE FUZZY LOGIC ALGORITHM

The fuzzy logic algorithm is more complex than the weighted average algorithm, yet still simple enough to be used easily. The algorithm has two branches to allow for different levels of user involvement. The fuzzy set membership functions used in the algorithm can be defined in two ways: automatic and manual. In the automatic method, the membership functions are defined based on the number of sets, the mean, the standard deviation, and a constant that is multiplied by the standard deviation. The automatic method is divided into two categories, triangular and trapezoidal. With the manual method, the user can control all parameters by defining any combination of shapes for the membership functions. In both methods, the weights for the membership functions and the modification factors are defined by the user.

Automatic Method

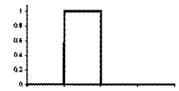
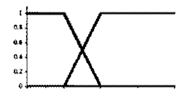
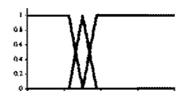
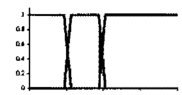
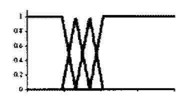
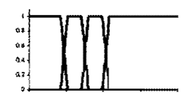
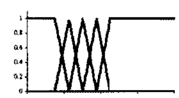
The automatic method depends on the generation of fuzzy sets for a given number of inputs. The automatic fuzzy sets are controlled by two statistical parameters, the mean and standard deviation of all the measurements for a particular parameter, and three user controlled inputs.

The first and second user inputs are the number of sets and set shapes. These two inputs go hand-in-hand and will be discussed together. For data sets that only some distinction between the high and low values or average and deviant values is needed, a small number of sets is acceptable. For data sets that are more critical and that require more subdivisions, up to five sets can be selected. When one set is selected, the distribution is rectangular, that is all of the degrees of membership are uniform. This results in a normal average, but some data can still be omitted where the function equals zero. For two sets, a z-shape and an s-shape function are used, as they complement each other. For three to five sets, combinations of z-shapes and s-shapes are used with triangular or trapezoidal shapes. Table 5.12 shows the possibilities for the automatic option based on the number of sets. The basic premise behind generating these sets is that the average of the data is directly in the center of the graph and the limits to either side of the center determined as a multiple of the standard deviation.

The third input is the spread in each set that is generated. These spreads are primarily influential when three or more sets are selected. The spread of the sets is controlled by multiplying a constant times the standard deviation. Figure 5.3 shows how changing the standard deviation multiplier changes the width of the graph. Only one example will be shown to explain how this multiplier works. Note how the center of the graph remains the same, but the sets are now wider. The general form for the limits of the sets is $\mu + f(n, shape)A\sigma$, where $f(n, shape)$ is a function that depends on the number of sets and shape of the sets and A is the width parameter. Tables 5.13 through 5.15 show the default limits for the automatic set generation. A is the multiplier for the width. The other constants were chosen in a manner to create symmetrical sets about the mean. The limits selected in Table 5.12 are defined by the equations shown in Tables 5.13 through 5.15. The boundaries are designated by where there is a corner on a continuous line, that is, the point of a triangle, the corner of a trapezoid or rectangle, or the bend in a z-shape or s-shape. This also includes where the triangle, rectangle, or trapezoidal shapes go to zero. The last two limits in each column will overlap with the first two in the following column. This is to ensure that the sum of all overlapping membership functions adds up to one, as wherever membership functions overlap, their sum must be equal to one. The sets are set up so that the width of either one triangle or trapezoid is one standard deviation multiplied by the width factor, that is, $width = A\sigma$. The z-shape or s-shape "ramps" on either side simply depend on the outer sets and the width is not applicable there. However in the case of two sets, where only z-shape and s-shape curves are used, the middle zone where the "X" is formed by the crossing graphs has a width of one standard deviation times the width factor. To interpret the numbers shown in Tables 5.13 through 5.15 match the number of sets and the set shape to the figures shown in

Table 5.12. For example for the 3-set triangle shape in Table 5.14, the middle three values describe where the triangle in Table 5.12 starts to increase, peaks, and then goes to zero, respectively. Figure 5.4 shows how this interpretation works. The other set shapes are defined in a similar manner, from left to right by set, and the from left to right for each bend or corner for that set shape.

Table 5.12 - Set Shapes for Automatic Set Generation

Number of Sets	Set Shapes Triangular/Trapezoidal	
1		
2		
3		
4		
5		

a) Low Multiplier

b) High Multiplier

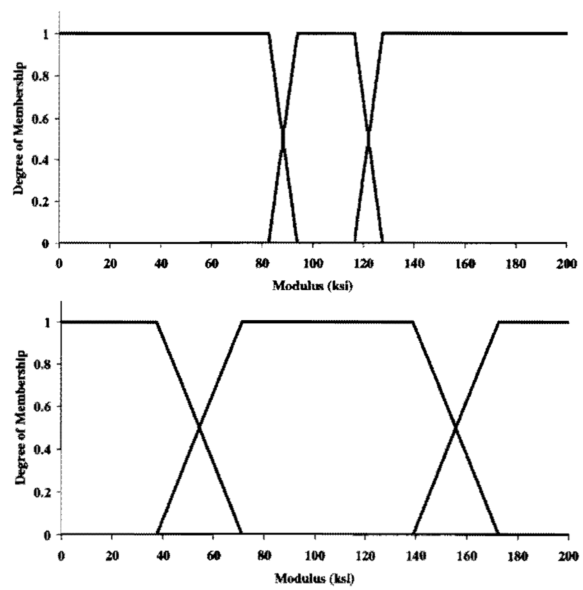


Figure 5.3 - Effect of Standard Deviation Multiplier

Table 5.13 - Set Limits for One and Two Sets

1 Set	
Rectangle	
$\mu - A\sigma$ $\mu + A\sigma$	
2 Sets	
Z-shape	S-shape
$\mu - A\sigma$ $\mu + A\sigma$	$\mu - A\sigma$ $\mu + A\sigma$

Table 5.14 - Set Limits for Three to Five Triangular Sets

3 Sets Triangular				
Z-shape	Triangle	S-shape		
$\mu - \frac{1}{2} A\sigma$ μ	$\mu - \frac{1}{2} A\sigma$ μ $\mu + \frac{1}{2} A\sigma$	μ $\mu + \frac{1}{2} A\sigma$		
4 Sets Triangular				
Z-shape	Triangle	Triangle	S-shape	
$\mu - \frac{3}{4} A\sigma$ $\mu - \frac{1}{4} A\sigma$	$\mu - \frac{3}{4} A\sigma$ $\mu - \frac{1}{4} A\sigma$ $\mu + \frac{1}{4} A\sigma$	$\mu - \frac{1}{4} A\sigma$ $\mu + \frac{1}{4} A\sigma$ $\mu + \frac{3}{4} A\sigma$	$\mu + \frac{1}{4} A\sigma$ $\mu + \frac{3}{4} A\sigma$	
5 Sets Triangular				
Z-shape	Triangle	Triangle	Triangle	S-shape
$\mu - A\sigma$ $\mu - \frac{1}{2} A\sigma$	$\mu - A\sigma$ $\mu - \frac{1}{2} A\sigma$ μ	$\mu - \frac{1}{2} A\sigma$ μ $\mu + \frac{1}{2} A\sigma$	μ $\mu + \frac{1}{2} A\sigma$ $\mu + A\sigma$	$\mu + \frac{1}{2} A\sigma$ $\mu + A\sigma$

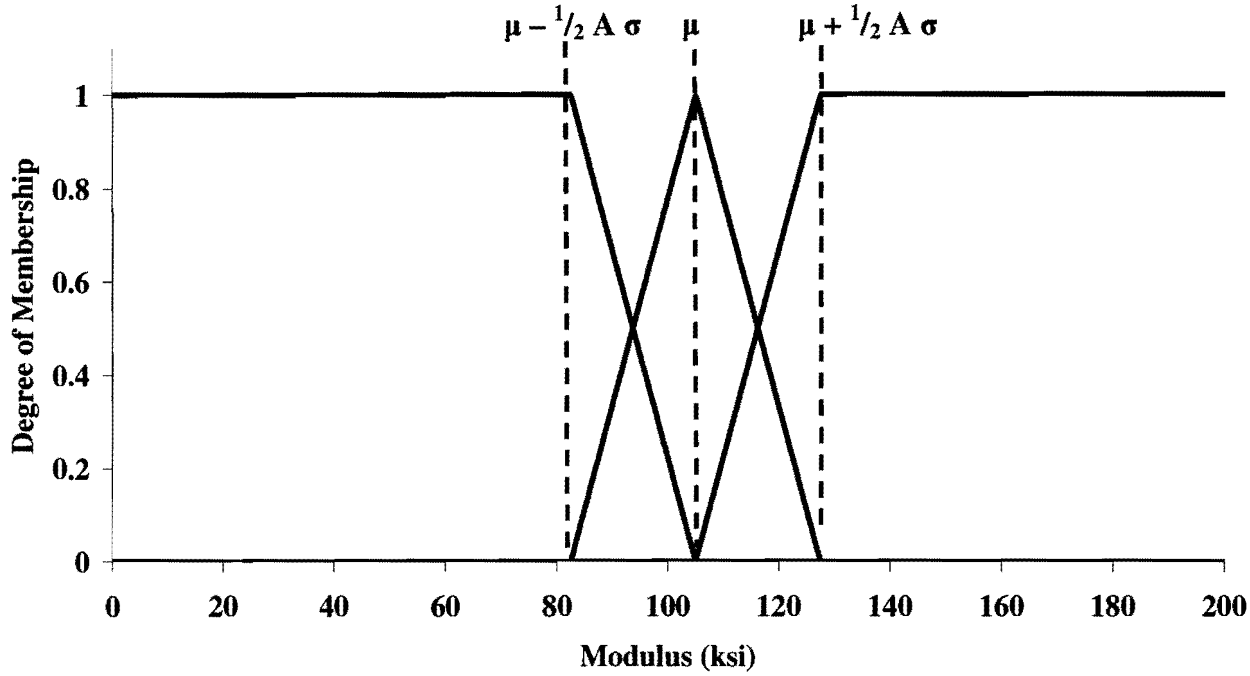


Figure 5.4 - Interpretation of Set Limits

The use of device modification factors will require a weighted mean and standard deviation to be calculated as discussed previously. The weighted standard deviation involved can be represented as

$$\sigma_{weighted} = \sqrt{\frac{\sum_{i=1}^{n} w_i (x_i - \mu_w)^2}{\frac{(n'-1)\sum_{i=1}^{n} w_i}{n'}}} \tag{5.3}$$

where μ_w is the weighted mean, x_i's the individual values, w_i's are the individual weights, n the number of data points, and the n' the number of non-zero weights (NIST, 2001). The weight used in Equation 5.3 is the device modification value. The example previously shown that fused the TxDOT data using fuzzy logic did not include the weighted average or weighted standard deviation. The purpose of the example was not to incorporate those concepts at that time. However, the weighted mean and weighted standard deviation are included in the algorithm for the generation of the automatic sets. The discussion on how to select appropriate device modification values is included later in this chapter.

Table 5.15 - Set limits for Three to Five Trapezoidal Sets

3 Sets Trapezoidal				
Z-shape	Trapezoid	S-shape		
$\mu - \tfrac{1}{2} A \sigma$ $\mu - \tfrac{1}{4} A \sigma$	$\mu - \tfrac{1}{2} A \sigma$ $\mu - \tfrac{1}{4} A \sigma$ $\mu + \tfrac{1}{4} A \sigma$ $\mu + \tfrac{1}{2} A \sigma$	$\mu + \tfrac{1}{4} A \sigma$ $\mu + \tfrac{1}{2} A \sigma$		
4 Sets Trapezoidal				
Z-shape	Trapezoid	Trapezoid	S-shape	
$\mu - \tfrac{7}{8} A \sigma$ $\mu - \tfrac{5}{8} A \sigma$	$\mu - \tfrac{7}{8} A \sigma$ $\mu - \tfrac{5}{8} A \sigma$ $\mu - \tfrac{1}{8} A \sigma$ $\mu + \tfrac{1}{8} A \sigma$	$\mu - \tfrac{1}{8} A \sigma$ $\mu + \tfrac{1}{8} A \sigma$ $\mu + \tfrac{5}{8} A \sigma$ $\mu + \tfrac{7}{8} A \sigma$	$\mu + \tfrac{5}{8} A \sigma$ $\mu + \tfrac{7}{8} A \sigma$	
5 Sets Trapezoidal				
Z-shape	Trapezoid	Trapezoid	Trapezoid	S-shape
$\mu - \tfrac{5}{4} A \sigma$ $\mu - A \sigma$	$\mu - \tfrac{5}{4} A \sigma$ $\mu - A \sigma$ $\mu - \tfrac{1}{2} A \sigma$ $\mu - \tfrac{1}{4} A \sigma$	$\mu - \tfrac{1}{2} A \sigma$ $\mu - \tfrac{1}{4} A \sigma$ $\mu + \tfrac{1}{4} A \sigma$ $\mu + \tfrac{1}{2} A \sigma$	$\mu + \tfrac{1}{4} A \sigma$ $\mu + \tfrac{1}{2} A \sigma$ $\mu + A \sigma$ $\mu + \tfrac{5}{4} A \sigma$	$\mu + A \sigma$ $\mu + \tfrac{5}{4} A \sigma$

Manual Method

In the manual set generation, any combination of the five shapes that the algorithm allows can be used. The user inputs the number of sets and the shape of each set. The user then defines the set boundaries. The width factor that was used with the automatic method is not used in the manual method. The sets on the left and right ends (the outer sets) are normally z-shaped and s-shaped, respectively. The outer sets can be any other shape, namely the triangular and trapezoidal shapes.

Assigning Weights to the Fuzzy Sets

Now that the differences between the two set generation methods have been discussed, some of the considerations that must be made when assigning weights to the fuzzy sets should be explained.

When a set of numbers (or moduli) exists for a given site, some idea of the expected site variability can easily be determined. Additionally, the user may have some idea of which moduli are unacceptable or which ones demand more attention. Typically, the average values are considered as the norm and should be weighted more heavily. If a high-end estimate is needed, values above the average should be weighted more. On the contrary, if there is much concern about low moduli, the low-end values should be weighted heavier to obtain a more conservative estimate. The weights assigned in the fuzzy logic example earlier in this chapter are shown in Table 5.16. The weights for each layer are relative to each other. If the weights for the AC layer are based on a scale of zero to one hundred and the weights of the base layer on zero to twenty, only the relative differences (ratio-wise) will impact the weighting of the inputs. When assigning weights to sets, the following factors should be considered: the significance of extreme values, the natural spread in the data, and the number and size of the sets. More sets will allow for more refinement in the weights. Larger sets could be used where only extreme values need to be removed and where values in the average range can be considered equally. Precise data sets may not require many sets, whereas the opposite may be true for more spread out data.

Table 5.16 - Assigned Weights for Fuzzy Sets

Layer	Fuzzy Set Weights			
	Far Below Average	**Below Average**	**Average**	**Above Average**
AC	1	2	3	2.5
Base	1.5	2	3	2.5
Subgrade	1.5	2	2.5	2.25

From the weights assigned to each layer, it can quickly be observed which layers will be impacted the most by high, low, or average data (magnitude-wise). For the AC layer, the below average data is not weighted heavily, as the weight for that set is one-third that of the average set weight. However, for the base and subgrade layers, the below average data are weighted heavier. The ratios 1.5:3 and 1.5:2.5 each progressively weigh below average data more. In those cases, the fused results are more conservative. It would also be possible to have a set that precedes Set 1 as shown in Table 5.16 for which the weight can be very low, for example, 0.1. This would mean that the below average data will be considered, but extremely outlying small values will be omitted.

Consider the following set of data: {20, 200, 240, 250, 150, 260, 300}. The value of 150 could still be considered to be a viable possibility as it is still fairly close to the rest of the data. For a conservative estimate, the 150 may be weighted heavier. The value of 20 must be scrutinized, since it is so far from the mean and could be completely out of the range of a feasible value. Thus the 20 should be omitted or weighted very low. In this case, the value is probably an error and should manually be removed if the end set has a high enough weight that the value of 20 will significantly alter the fused value. Assume that the weights are defined by the following set: {0, 2.5, 1.75, 2, 1.5}. This would provide a conservative estimate up to a point. Values beyond that point will be omitted (represented by the zero).

The weight value is calculated as a weighted average based on the degree of membership (refer to Chapter 3) and the associated factor for each set. The assigned weights with the degrees of membership for each value are combined to obtain the weight to be used to fuse the data based on

$$w_i = \frac{\sum D_m w_{mf}}{\sum D_m} \tag{5.4}$$

The D_m's refer to the degrees of membership and the w_{mf}'s represent the assigned weight of the function (e.g. 1, 2, 3).

Defining of Device Modification Factors

In the event that expert opinion or additional information is available about the reliability of a device or a particular layer for a device, the additional modification factors can be added. These modification factors can be based on the overall reliability of a particular device, the reliability of the operator using the device, or the expert opinion of the user on the quality of the data or results. The input for this part of the procedure will be in the form of a decimal number between zero and one that represents the modification to be applied to the data. As long as the number is greater than zero, the values for that parameter will still be included when the average and standard deviation values are taken. Only when a set of values for a particular parameter for a device is completely omitted (i.e. the modification value is zero), will they not be included in the calculation of the mean and standard deviation. This was done because there are times where the modulus of a layer might be fixed and not backcalculated at all. This is where a modification value of zero may be used. However, if the fixed modulus value of the layer seems credible, then a number greater than zero might be used to describe the degree of belief. For example, if previous data exists that suggests an average value for a site and if the reliability is fifty percent, then the modification value for that list of numbers will be 0.5 and the numbers will be included in the calculation of the mean and standard deviation of the data.

Once all inputs have been entered, the calculation is performed based on the mathematical procedure described at the beginning of this chapter. The flowchart in Figure 5.5 shows at which point in the reduction procedure each input is needed and where the automatic and manual methods converge.

After the required inputs are entered, they can be adjusted based on any criteria that the user may want to include. For example, the user may want to see how the results change if data from a device is completely or partially not included in the fusing of the data.

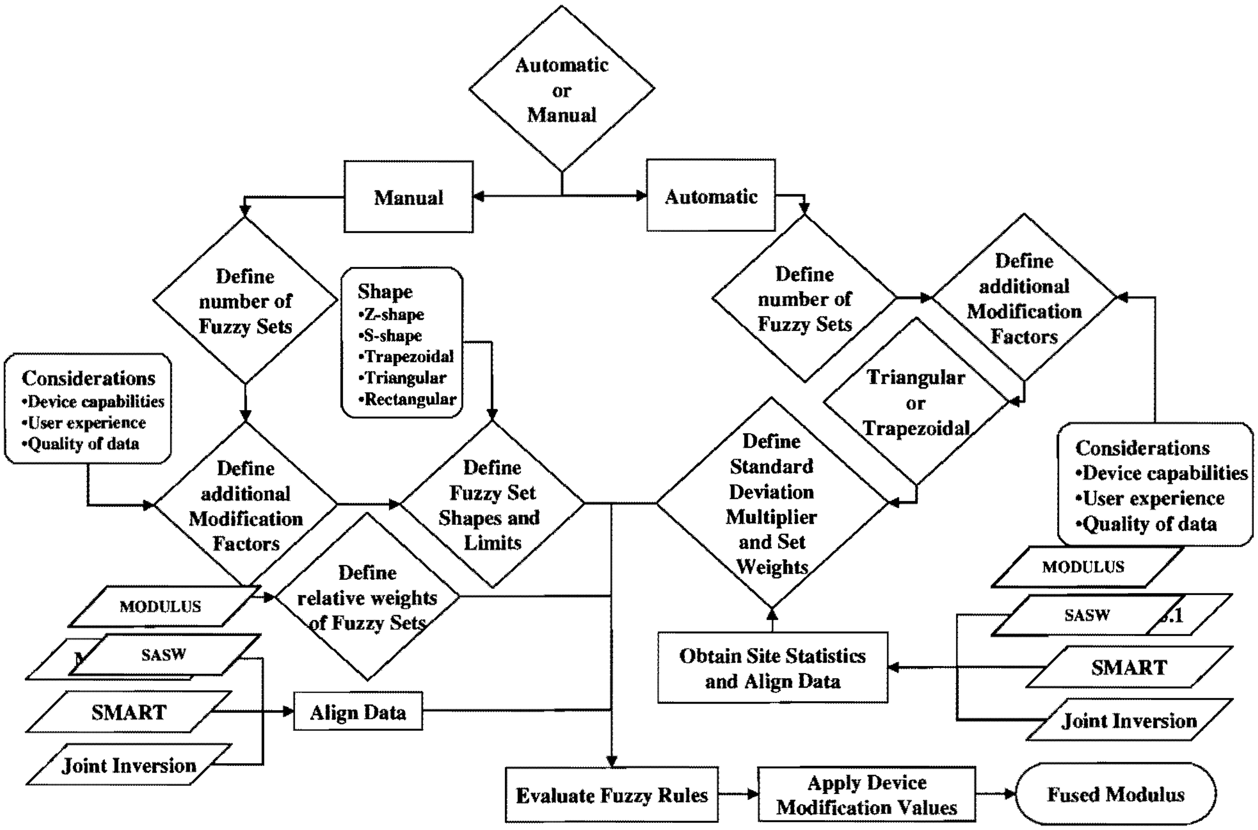


Figure 5.5 - Fuzzy Logic Algorithm

CHAPTER SIX

HYBRID METHOD

INTRODUCTION

The hybrid method combines aspects of the fuzzy logic method and the weighted average method. The basic premise of this method is to use the weighted average method equations to combine the data, but to modify the input using fuzzy logic.

Recall that with the weighted average method, problems can arise where discrepant values exist. Additionally, lower values would always have a higher weight as a constant coefficient of variation was always used. Both of these two potential problems can be solved by modifying the input by means of a modification factor that is obtained from fuzzy logic. The input is modified by applying modification factors to the coefficients of variation before the fusion calculation using the weighted average method. These modification factors are based on the deviation of moduli from the mean and can also be based on whether the moduli are above or below the mean. The deviation of moduli from the mean can be used to eliminate outlying data. By modifying values above and below the mean slightly differently, two sets of numbers with similar coefficients of variation will not have higher values weighted less due to the magnitude of the modulus being greater.

DEVELOPMENT OF HYBRID METHOD

Consider a situation where there are two distributions defined as $\mathcal{N}(100, 20)$ and $\mathcal{N}(120, 24)$. Both distributions have a coefficient of variation of 20%. Additionally, assume that the site mean and standard deviation are 110 and 28 (includes both sets of data as one). Now consider two values that are to be combined, one from each set with values of 110 and 105. The value of 105 will have a higher weight as the weight is proportional to the inverse of the product of the value and the coefficient of variation. Although the value of 110 is at the site average and less deviant from its device's average than the 105 value, it will be weighted less because its absolute value is greater than the other value (105). This is where the fuzzy logic method can be employed to alter the weights such that the values that are deviant from the mean are considered less. Additionally, a modification can also be made so that values of equal distance both above and below the mean will have more similar weights. The weights will be fairly close to begin with if the coefficients of variation are close. This difference in the weights is usually small enough to ignore if the two values are fairly close, that is $22^{-2} \approx 21^{-2}$ (weights are proportional to the inverse of the standard deviation squared). However, if one value is significantly lower than the other, but obviously out of the expected range, it will dominate the combining of the values, e.g. $10^{-2} >> 21^{-2}$.

Defining Modification Values

The modification values are found using a fuzzy logic procedure. In the fuzzy logic method, the results of taking the weighted average of the triggered membership functions were the weights used to take the weighted average of the moduli. For the hybrid method, the fuzzy logic procedure produces a weighted modification value based on fuzzy sets. The outputs of applying the fuzzy sets and membership functions along with the set weights will instead be a factor between zero and one for which a composite modification factor will be found using a weighted average. The hybrid modification factor, M_H, is a function of the degree of membership, D_m, and the associated factor, w_f, for each fuzzy set, using.

$$M_H = \frac{\sum D_m w_f}{\sum D_m} \tag{6.1}$$

Equation 6.1 is essentially the same as Equation 5.3; the only difference is the nomenclature.

The w_f's, or the value assigned as the output for when a fuzzy rule is triggered, must be defined so that the coefficients of variation used in the statistical weighted average part of the algorithm are modified properly. It is important to point out that the modification factors for the coefficients of variation are divided (not multiplied) for this algorithm. Hence, values lower than one will increase the coefficient of variation and a value of one will not change it. Values above one are not recommended as that would make the coefficient of variation smaller. It is assumed that the theoretical (or unmodified) coefficient of variation is the smallest value that should be used. The idea behind modifying the coefficients of variation is to penalize data that is deviant and leave data that is closer to the mean intact by not modifying it (or modifying it slightly). Therefore, in assigning the weights for the sets, a procedure similar to the fuzzy logic algorithm can be implemented. Figure 6.1 shows one option for assigning weights for a given set of data. This is the same data that was fused in Chapters 4 and 5.

Assume that the center triangle is at the average and that values to the left represent numbers smaller than the average and values to the right are greater than the average. In Figure 6.1, the values smaller than the average require more modification, that is, the coefficient of variation will be divided by a smaller number resulting in a greater change. The altered coefficients of variation will always be greater than the original. This will represent greater uncertainty in values that are deviant from the mean.

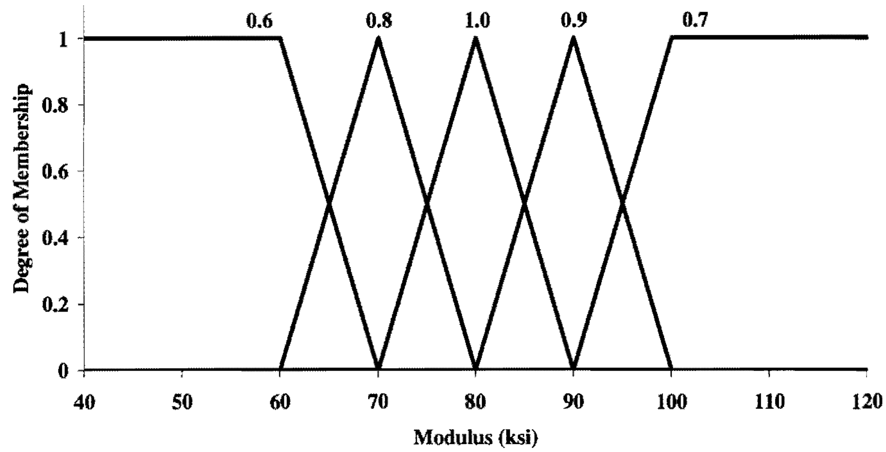


Figure 6.1 - Example of Weights for Hybrid Algorithm

To illustrate the application of this method, the data that was shown in Table 4.1 will once again be considered. The coefficients of variation are repeated in Table 6.1. The 1000000% coefficient of variation for the AC modulus of MODULUS procedure will still be used. Recall that this coefficient of variation was applied because the modulus of the AC layer was fixed, not calculated.

Table 6.1 - Coefficients of Variation for Hybrid Method

Layer	Coefficient of Variation		
	MODULUS	SMART	JIM
AC	1000000%	14%	14%
Base	46%	42%	18%
Subgrade	10%	16%	11%

The values from Table 4.1 will be used to obtain the values which will modify their coefficients of variation. The same fuzzy sets will be used that were used in the previous chapter, except that the weights of the fuzzy sets are different. Refer to Table 5.2 and Figure 5.1 for the set boundaries. There were four sets defined for each parameter. Each set must have an output associated with it. This output is the weight. More specifically, this weight is the modification factor for the coefficient of variation. Once the fuzzy procedure is performed for each input, a composite coefficient of variation modification factor can be obtained by taking a simple weighted average. For this example, the weights shown in Table 6.2 are assigned to each parameter. The sets are ordered from left to right as they were in Chapter 5.

Each point has three moduli associated with it for each layer. As was done with the fuzzy logic procedure, the moduli were evaluated based on the fuzzy sets that were presented in Figure 5.1. The fuzzy membership values are obtained for each point and combined using Equation 6.1. The modification factors for the coefficients of variation are shown in Table 6.3 for each test point.

The numbers at the top of each column to the right of the row labeled "CoV" are the site coefficients of variation. They are the unmodified coefficients of variation. The numbers to the right of the numbered rows are the modification factors that will be applied to the coefficients of variation shown in the "CoV" row. Each modification factor is applied to the coefficient of variation that is in the same column.

Table 6.2 - Coefficient of Variation Modification Values per Set

Layer	CoV Modification Factor			
	Far Below Average	Below Average	Average	Above Average
AC	0.3	0.5	1	0.7
Base	0.5	1	.9	.6
Subgrade	0.9	1	0.8	0.7

Table 6.3 - Modification Values for Coefficients of Variation

Test Pt.	Modification Factor for Coefficient of Variation								
	AC			Base			Subgrade		
	MODULUS	SMART	JIM	MODULUS	SMART	JIM	MODULUS	SMART	JIM
CoV	10^6%	14%	14%	46%	42%	18%	10%	16%	11%
1	0.67	0.70	0.70	0.98	0.80	1.00	0.70	0.70	0.92
2	0.67	0.70	0.70	0.79	1.00	0.98	0.70	0.70	0.92
3	0.67	0.70	0.70	0.92	0.66	0.92	0.73	0.76	0.96
4	0.67	0.98	0.67	0.94	0.81	0.97	0.75	0.70	1.00
5	0.67	0.64	0.98	0.98	0.75	0.97	0.73	0.70	1.00
6	0.67	0.70	0.78	0.75	0.60	0.75	0.72	0.70	1.00
7	0.67	0.70	0.70	0.60	0.60	0.92	0.71	0.70	0.96
8	0.67	0.72	0.99	0.60	0.60	0.92	0.72	0.80	0.96
9	0.67	0.70	0.70	0.60	0.60	0.92	0.75	0.88	1.00
10	0.67	0.84	0.72	0.93	0.75	0.95	0.79	0.76	0.98
11	0.67	0.31	0.33	0.79	0.95	0.99	0.82	0.78	0.96
12	0.67	0.72	0.86	0.95	0.60	0.90	0.78	0.70	0.98
13	0.67	0.92	0.85	0.75	0.75	0.95	0.74	0.70	1.00
14	0.67	0.45	0.60	0.96	0.60	0.92	0.75	0.84	1.00
15	0.67	0.76	0.80	0.75	0.75	0.75	0.76	0.84	1.00
16	0.67	0.81	0.57	0.99	0.75	0.92	0.76	0.70	0.98
17	0.67	0.95	0.85	0.96	0.80	0.97	0.77	0.70	0.98
18	0.67	0.99	0.84	0.99	0.96	0.98	0.77	0.76	0.98
19	0.67	0.95	0.79	0.84	0.77	0.99	0.78	0.76	0.98
20	0.67	0.70	0.70	0.75	0.75	0.80	0.74	0.84	1.00

Next, each coefficient of variation is divided by the corresponding modification factors. The result is increased coefficients of variation. Table 6.4 shows the results of this computation.

Table 6.4 - Modified Coefficient of Variation Values

Test Pt.	Modified Coefficient of Variation								
	AC			Base			Subgrade		
	MODULUS	SMART	JIM	MODULUS	SMART	JIM	MODULUS	SMART	JIM
CoV	10^6%	14%	14%	46%	42%	18%	10%	16%	11%
1	15E5%	20%	20%	47%	52%	18%	15%	23%	11%
2	15E5%	20%	20%	59%	42%	18%	15%	23%	11%
3	15E5%	20%	20%	50%	63%	19%	14%	21%	11%
4	15E5%	14%	20%	49%	51%	18%	14%	23%	11%
5	15E5%	22%	14%	48%	56%	18%	14%	23%	11%
6	15E5%	20%	18%	61%	69%	23%	14%	23%	11%
7	15E5%	20%	20%	77%	69%	19%	15%	23%	11%
8	15E5%	20%	14%	77%	69%	19%	14%	20%	11%
9	15E5%	20%	20%	77%	69%	19%	14%	18%	11%
10	15E5%	17%	19%	50%	56%	18%	13%	21%	11%
11	15E5%	45%	41%	58%	44%	18%	13%	21%	11%
12	15E5%	20%	16%	48%	69%	19%	13%	23%	11%
13	15E5%	15%	16%	61%	56%	19%	14%	23%	11%
14	15E5%	31%	23%	48%	69%	19%	14%	19%	11%
15	15E5%	18%	17%	61%	56%	23%	14%	19%	11%
16	15E5%	17%	24%	46%	56%	19%	14%	23%	11%
17	15E5%	15%	16%	48%	52%	18%	14%	23%	11%
18	15E5%	14%	16%	46%	43%	18%	14%	21%	11%
19	15E5%	15%	17%	55%	54%	18%	13%	21%	11%
20	15E5%	20%	20%	61%	56%	22%	14%	19%	11%

After the new coefficients of variation have been found, the remainder of the procedure is identical to the weighted average method. Equations 2.6 through 2.8 are used from this point on. The results of this procedure are shown in Table 6.5. If Table 6.5 is compared to Table 4.3, it is evident that while similar in result, the variation reported in Table 6.5 is higher. This was due to the modification of the coefficients of variation.

Table 6.5 - Results of Hybrid Method

Test Pt.	AC Modulus (*ksi*)			Base Modulus (*ksi*)			Subgrade Modulus (*ksi*)		
	Avg	StDev	CoV	Avg	StDev	CoV	Avg	StDev	CoV
1	659	93	14%	62	10	16%	14	1	9%
2	634	89	14%	63	10	16%	14	1	9%
3	644	90	14%	82	14	18%	13	1	8%
4	535	63	12%	64	11	16%	12	1	8%
5	530	63	12%	74	12	16%	12	1	8%
6	626	83	13%	109	23	21%	12	1	8%
7	704	99	14%	96	17	18%	13	1	9%
8	534	60	11%	94	17	18%	13	1	8%
9	711	100	14%	95	17	18%	11	1	8%
10	602	76	13%	83	14	17%	11	1	8%
11	408	125	31%	62	10	16%	10	1	8%
12	598	74	12%	96	17	18%	11	1	8%
13	549	61	11%	85	14	17%	12	1	8%
14	477	88	19%	91	16	17%	12	1	8%
15	605	76	13%	107	22	20%	12	1	8%
16	508	72	14%	86	15	17%	11	1	8%
17	545	60	11%	66	11	16%	11	1	8%
18	540	58	11%	66	10	16%	11	1	8%
19	565	64	11%	63	10	16%	11	1	8%
20	661	93	14%	104	20	19%	12	1	8%

The hybrid method allows for moduli farther from the mean (or any selected datum) to be weighted less, or possibly more, depending on the weights that are assigned. To make an average more conservative the greater modification factors should be placed on the left side membership sets whereas for a higher-end estimate, the numbers to the right side membership sets should be higher. Again, this factor should be limited to no more than one, where one represents no change in the coefficient of modification. Furthermore, had the values been highly spread, meaning that there were some extremely high or extremely low outlying data points, by setting a lower value for the weights on either end, the modified coefficient of variation will be much greater and subsequently weighted much less.

CHAPTER SEVEN

BAYESIAN INFERENCE METHOD

INTRODUCTION

In Chapter 3 the Bayesian inference method was introduced. In this chapter, that procedure will be applied to the NDT data from the TxDOT parking lot. The method is fairly straightforward. Defining the prior probabilities and conditional densities is the only complication.

DEFINING OF CONDITIONAL DENSITIES AND POSTERIOR PROBABILITIES

The conditional densities will be found by using a normal distribution for each method of data reduction. This distribution will be found by simply taking the average and standard deviation for each parameter for each method. Again that will be three methods, each providing values for three layers for a total of nine distributions. Naturally, only data from like layers will be combined.

Equations 2.13 and 2.14 will be employed. Both of the options for calculating the prior probabilities, namely, the cumulative approach and the double-tailed approach will be used. The AC modulus from MODULUS will have a coefficient of variation of 50%, instead of 1000000%. Tables 7.1 and 7.2 show the results of applying the Bayesian method to the TxDOT data.

INTERPRETATION OF POSTERIOR PROBABILITIES

The outputs shown in Tables 7.1 and 7.2 can be interpreted in many ways. One is to make a hard decision based on the posterior probabilities and to select the modulus value that corresponds to the value with the highest posterior probability. For example, for Test Point 2 in Table 7.2 for the AC modulus, the SMART modulus will be selected as representative for that test point as its posterior probability is the highest of the three inputs at that point. For Test Point 5 the JIM result would be selected. The results for using this decision method for obtaining results using the Bayesian method are shown in Table 7.3.

Table 7.1 - Posterior Probabilities for Cumulative Method

Test Pt.	Posterior Probability								
	AC			Base			Subgrade		
	M†	S†	J†	M	S	J	M	S	J
1	14%	48%	39%	47%	45%	8%	18%	49%	33%
2	12%	44%	44%	20%	13%	67%	8%	55%	37%
3	12%	45%	42%	6%	21%	73%	35%	13%	52%
4	27%	57%	16%	22%	2%	76%	25%	19%	55%
5	29%	15%	56%	31%	27%	42%	31%	20%	49%
6	13%	41%	47%	33%	20%	48%	33%	17%	50%
7	21%	24%	55%	12%	18%	70%	32%	18%	51%
8	27%	20%	54%	12%	2%	85%	39%	4%	58%
9	23%	55%	22%	2%	17%	80%	30%	0%	70%
10	12%	40%	47%	24%	22%	54%	15%	43%	42%
11	98%	1%	1%	24%	40%	36%	5%	88%	7%
12	13%	47%	41%	18%	22%	61%	21%	47%	32%
13	20%	54%	26%	27%	16%	57%	29%	21%	50%
14	65%	7%	28%	15%	18%	68%	30%	2%	69%
15	12%	44%	44%	31%	22%	48%	25%	2%	73%
16	41%	43%	15%	9%	18%	73%	34%	40%	27%
17	21%	52%	27%	45%	2%	53%	30%	42%	28%
18	24%	47%	29%	37%	7%	56%	36%	32%	32%
19	16%	27%	58%	20%	57%	23%	16%	42%	41%
20	14%	41%	45%	30%	16%	54%	37%	1%	62%

† M=MODULUS, S = SMART, J = JIM

Table 7.2 - Posterior Probabilities for Double-Tailed Method

Test Pt.	Posterior Probability								
	AC			Base			Subgrade		
	M†	S†	J†	M	S	J	M	S	J
1	42%	41%	17%	47%	45%	8%	3%	92%	5%
2	28%	39%	33%	20%	13%	67%	0%	95%	5%
3	33%	42%	25%	13%	26%	61%	40%	34%	26%
4	27%	57%	16%	22%	2%	76%	32%	4%	64%
5	29%	15%	56%	31%	27%	42%	27%	9%	64%
6	25%	18%	57%	61%	17%	22%	19%	5%	76%
7	75%	3%	22%	3%	14%	84%	38%	15%	46%
8	27%	20%	54%	2%	0%	98%	49%	13%	38%
9	78%	20%	2%	0%	8%	92%	32%	0%	68%
10	18%	51%	31%	25%	17%	58%	15%	43%	42%
11	98%	1%	1%	24%	40%	36%	5%	88%	7%
12	17%	30%	52%	46%	27%	26%	35%	14%	52%
13	20%	54%	26%	13%	19%	68%	32%	9%	59%
14	65%	7%	28%	33%	9%	58%	31%	2%	67%
15	18%	39%	43%	37%	52%	11%	27%	2%	71%
16	41%	43%	15%	17%	30%	52%	50%	10%	40%
17	21%	52%	27%	45%	2%	53%	45%	11%	43%
18	24%	47%	29%	37%	7%	56%	36%	32%	32%
19	20%	34%	46%	20%	57%	23%	16%	42%	41%
20	45%	24%	32%	38%	45%	16%	34%	2%	65%

† M=MODULUS, S = SMART, J = JIM

Table 7.3 - Modulus Values from Maximum Posterior Bayesian Method

Test Pt.	Modulus (*ksi*)					
	AC		Base		Subgrade	
	Cumulative	Double-tailed	Cumulative	Double-tailed	Cumulative	Double-tailed
1	647	500	68	68	20	20
2	636	632	68	68	20	20
3	636	636	92	92	11	18.7
4	556	556	73	73	10	10
5	547	547	72	72	10	10
6	605	605	105	119	10	10
7	681	500	93	93	11	11
8	549	549	91	91	11	18.9
9	689	500	92	92	10	10
10	619	590	80	80	17	17
11	500	500	81	81	16	16
12	621	584	99	78	21	9
13	571	571	82	82	10	10
14	500	500	92	92	10	10
15	601	601	104	127	10	10
16	521	521	92	92	21	16.8
17	563	563	72	72	21	16.6
18	552	552	69	69	16.7	16.7
19	603	603	103	103	17	17
20	655	500	102	113	10	20

Another option for interpreting the results would be to use the posterior probabilities as weights and to simply take a weighted average. Hence, for Test Point 1 for the base layer in Table 7.1, 47% of the value from MODULUS, 45% from the SMART, and 8% from the JIM will be used. As the sum of the posterior probabilities is one, the numerator in the weighted average calculation essentially drops out of the calculation. The results from fusing the data using this weighted average approach with Bayesian-developed weights are shown in Table 7.4.

Table 7.4 - Modulus Values from Weighted Average Bayesian Method

Test Pt.	Modulus (*ksi*)					
	AC		Base		Subgrade	
	Cumulative	Double-tailed	Cumulative	Double-tailed	Cumulative	Double-tailed
1	637	589	83	83	18	20
2	618	597	62	62	17	20
3	626	595	99	99	14	16
4	532	532	68	68	14	13
5	526	526	82	82	15	13
6	614	588	125	126	15	12
7	654	547	114	105	15	16
8	528	528	103	92	14	15
9	655	542	109	99	12	12
10	593	583	94	91	13	13
11	498	498	66	66	15	15
12	591	581	106	102	16	13
13	545	545	96	93	14	13
14	494	494	103	94	12	12
15	592	586	111	118	12	12
16	507	507	95	97	16	14
17	540	540	74	74	16	14
18	532	532	65	65	14	14
19	571	562	83	83	14	14
20	639	589	105	109	13	13

DEFINING UNCERTAINTY FOR BAYESIAN FUSION

The uncertainty in the combined values can also be found by applying an equation similar to Equation 5.2. The only change is that there are no modification factors (k_m's), so Equation 5.2 simplifies to

$$\sigma_F = \sqrt{\frac{\sum (w_i \sigma_i)^2}{\left(\sum (w_i)\right)^2}} \tag{7.1}$$

This will represent the uncertainty in the results using the Bayesian process to fuse the moduli. Table 7.5 shows the uncertainties that will be obtained from using the cumulative and double-tailed approaches. If either of the Bayesian decision methods is used, the uncertainty is considered to be the same as the *a priori* uncertainty of the selected value and can be in the form of either standard deviation or coefficient of variation.

Table 7.5 - Uncertainty Values from Weighted Average Bayesian Method

Test Pt.	Uncertainty as Standard Deviation (*ksi*)					
	AC		Base		Subgrade	
	Cumulative	Double-tailed	Cumulative	Double-tailed	Cumulative	Double-tailed
1	66	113	24	24	2	3
2	62	83	10	10	2	3
3	64	93	17	18	1	1
4	81	81	11	11	1	1
5	85	85	17	17	1	1
6	63	81	25	36	1	1
7	78	189	19	17	1	1
8	79	79	16	16	1	1
9	82	196	18	16	1	1
10	61	67	17	16	1	1
11	245	245	15	15	2	2
12	61	66	18	24	2	1
13	69	69	18	15	1	1
14	164	164	17	16	1	1
15	60	66	21	33	1	1
16	109	109	15	19	1	1
17	70	70	17	17	2	1
18	72	72	12	12	1	1
19	65	68	25	25	1	1
20	66	118	19	29	1	1

CHAPTER EIGHT

PRESENTATION OF FUSED RESULTS

INTRODUCTION

Now that the fusion methods have been developed, a summary of fused data from two test sites will be given. The data from the TxDOT site that was used to explain the methods will again be fused. In addition, data from a second site, the Texas Transportation Institute (TTI) site will also be fused. Both sites had SPA and FWD available which would allow for SMART, MODULUS, and JIM analyses to be performed. In addition, the TTI site also had GPR and PSPA data available. The PSPA data for the top layer will be added to the fusion process. The GPR data is not being used since the other methods used a fixed thickness and did not backcalculate thickness and thus the fusion process would consider the GPR data as a reading and all other data as a best guess. The methods developed in the previous four chapters will be applied. It is essential to point out that the method that shows the most promise overall is the fuzzy logic method. It allows for the most flexibility and allows both for statistical and expert inputs in the fusion process.

TXDOT PARKING LOT REVISITED

Fuzzy Logic Fusion

First the TxDOT data is revisited using the fuzzy logic method. The three moduli of the layers (AC, base, and subgrade) will be rounded to the nearest 25, 15, and 5, respectively. All data will be fully weighted, meaning the device modification factors will be one, except for the AC modulus from MODULUS which will receive a weight of zero. The set boundaries are as shown in Table 8.1. The values shown in Table 8.1 represent the center of the fuzzy sets. As the interior sets are all symmetric triangles, the centers of the set correspond to the peak of the triangle. The edge sets are z on the left (low) and s on the right (high). The weights for the fuzzy sets are as represented in Table 8.2, and were arbitrarily chosen.

Table 8.1 - New Set Boundaries for TxDOT Parking Lot Data Revisited

Parameter	**Modulus (*ksi*)**		
	AC	**Base**	**SG**
Average	575	105	15
StDev	75	45	5
Set	Membership Boundaries (*ksi*)		
Very Low	425	15	5
Low	500	60	10
Average	575	105	15
High	650	150	20
Very High	725	195	25

Table 8.2 - Weights for Membership Functions Using Data from TxDOT Parking Lot

Layer	**Fuzzy Set Weights**				
	Very Low	**Low**	**Average**	**High**	**Very High**
AC	1	2	3	2.5	2
Base	1.5	2	3	2.5	1.5
Subgrade	1.5	2	2.5	2.25	2

The fuzzy set shapes are shown in Figure 8.1. The values used to plot these fuzzy sets were given in Table 8.1. The weights from Table 8.2 are not shown on the graphs. Set 1 is the left most set and Set 5 is the right most set. The device modification factors are as shown in Table 8.3.

Using the input shown in Tables 8.1 through 8.3 and the data from Table 4.1, the results of the fuzzy logic method are shown in Table 8.4. The uncertainty values are included with the fused values.

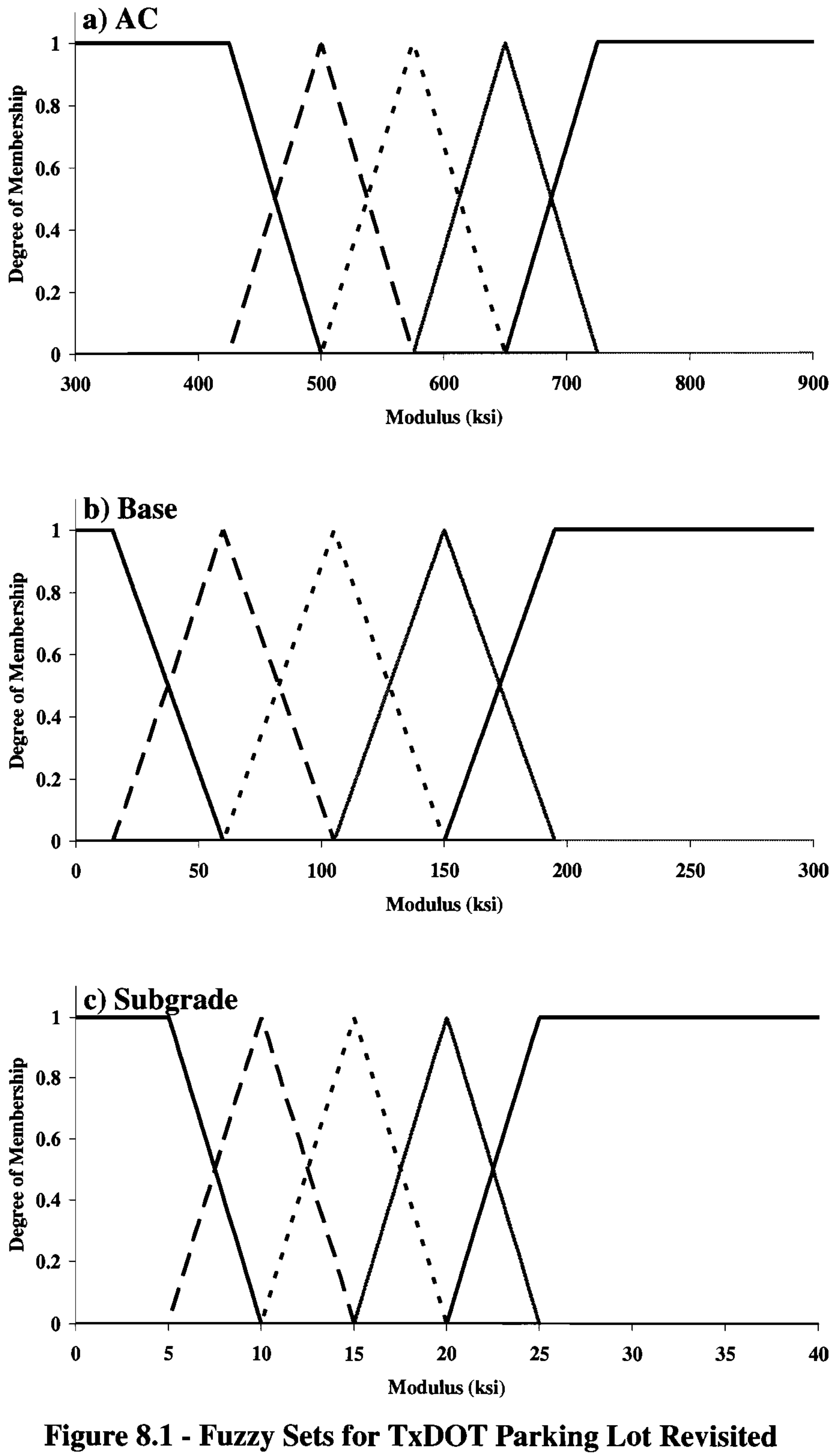


Figure 8.1 - Fuzzy Sets for TxDOT Parking Lot Revisited

Table 8.3 - Device Modification Values for Data from TxDOT Parking Lot

Method	Device Modification Factor		
	AC	Base	Subgrade
MODULUS	0	1	1
SMART	1	1	1
JIM	1	1	1

Table 8.4 - Results for TxDOT Data from Fuzzy Logic (Revisited) Using Data from TxDOT Parking Lot

Test Pt.	Modulus (*ksi*)								
	AC			Base			Subgrade		
	Avg	StDev	CoV	Avg	StDev	CoV	Avg	StDev	CoV
1	657	66	10%	82	22	27%	16	1	7%
2	634	62	10%	61	12	20%	17	1	7%
3	644	63	10%	106	25	24%	15	1	7%
4	534	54	10%	60	10	17%	15	1	8%
5	527	53	10%	89	22	25%	15	1	7%
6	629	62	10%	130	27	20%	15	1	7%
7	681	94	14%	131	29	22%	16	1	7%
8	532	53	10%	147	33	23%	14	1	7%
9	689	97	14%	137	30	22%	13	1	7%
10	604	59	10%	104	25	24%	13	1	7%
11	408	40	10%	68	16	24%	12	1	8%
12	602	59	10%	113	25	22%	14	1	8%
13	551	55	10%	103	24	23%	15	1	7%
14	475	47	10%	115	27	23%	13	1	7%
15	605	60	10%	114	24	21%	13	1	7%
16	506	51	10%	102	24	23%	14	1	8%
17	547	54	10%	64	12	18%	14	1	8%
18	540	54	10%	63	12	19%	13	1	7%
19	573	56	10%	81	22	27%	13	1	7%
20	661	65	10%	107	22	21%	13	1	7%
Avg	554	--	--	98	--	--	15	--	--
StDev	52	--	--	32	--	--	1.5	--	--
CoV	9%	--	--	32%	--	--	10%	--	--

Hybrid Fusion

The same fuzzy sets were used for the hybrid method that was used for the fuzzy logic method. In this case the outputs of the membership functions will be modification values for the coefficients of variation. The modification values are shown in Table 8.5. These values are represented by the weights of the fuzzy sets.

Table 8.5 - TxDOT Modification Values for Coefficients of Variation

Layer	Fuzzy Set Weights				
	Very Low	Low	Average	High	Very High
AC	0.75	0.85	1	0.9	0.8
Base	0.8	0.9	1	0.9	0.7
Subgrade	0.85	0.95	1	0.8	0.7

Using the inputs shown in Tables 8.1, 8.3 and 8.5 and the data from Table 4.1, the results form the hybrid method are shown in Table 8.6.

Bayesian Fusion

In revisiting the Bayesian method, the cumulative and double-tailed decision methods and the weighted average method were considered. The AC moduli from MODULUS are omitted. By applying the Bayesian methods to the data from Table 5.1, including the means and standard deviations, the results shown in Tables 8.7 through 8.9 are obtained. The decision method selects one modulus over another whereas the weighted average method uses the posterior probabilities as weights to obtain representative moduli.

Statistical Weighted Average Fusion

To be thorough, the statistical weighted average representative moduli for the TxDOT parking lot fusion are shown in Table 8.10.

Table 8.6 - Results for TxDOT Data from Hybrid Method Revisited

Test Pt.	Modulus (*ksi*)								
	AC			Base			Subgrade		
	Avg	StDev	CoV	Avg	StDev	CoV	Avg	StDev	CoV
1	659	73	11%	63	11	17%	14	1	8%
2	634	68	11%	62	11	17%	14	1	8%
3	644	70	11%	85	14	16%	13	1	8%
4	528	57	11%	64	11	17%	12	1	8%
5	522	57	11%	75	12	16%	12	1	8%
6	627	67	11%	109	17	16%	12	1	8%
7	702	83	12%	98	16	17%	13	1	8%
8	529	57	11%	95	16	17%	13	1	8%
9	709	85	12%	97	16	17%	12	1	7%
10	604	62	10%	84	14	16%	11	1	7%
11	408	54	13%	61	10	17%	10	1	7%
12	600	61	10%	98	16	16%	11	1	8%
13	549	57	10%	87	14	16%	12	1	8%
14	473	57	12%	92	15	16%	12	1	7%
15	605	62	10%	107	17	16%	12	1	7%
16	503	57	11%	88	14	16%	11	1	8%
17	545	57	10%	65	11	17%	11	1	8%
18	539	57	11%	66	11	17%	11	1	8%
19	571	59	10%	63	11	17%	11	1	7%
20	661	73	11%	104	17	16%	12	1	8%
Avg	580	--	--	99	--	--	12	--	--
StDev	73	--	--	27	--	--	1	--	--
CoV	13%	--	--	27%	--	--	8%	--	--

Table 8.7 - Modulus Values from Decision Bayesian Method Using TxDOT Parking Lot Results

Test Pt.	Modulus (*ksi*)					
	AC		Base		Subgrade	
	Cumulative	Double-tailed	Cumulative	Double-tailed	Cumulative	Double-tailed
1	647	647	68	68	20	20
2	636	632	68	68	20	20
3	636	636	92	92	11	18.7
4	556	556	73	73	10	10
5	547	547	72	72	10	10
6	605	605	105	118.9	10	10
7	681	681	93	93	11	11
8	549	549	91	91	11	18.9
9	689	689	92	92	10	10
10	619	590	80	80	17	17
11	412	412	81	81	16	16
12	621	584	99	78.3	21	9
13	571	571	82	82	10	10
14	490	490	92	92	10	10
15	601	601	104	127	10	10
16	521	521	92	92	21	16.8
17	563	563	72	72	21	16.6
18	552	552	69	69	16.7	16.7
19	603	603	103	103	17	17
20	655	655	102	113	10	10

Table 8.8 - Modulus Values from Weighted Bayesian Method Using TxDOT Parking Lot Results

Test Pt.	Modulus (*ksi*)					
	AC		Base		Subgrade	
	Cumulative	Double-tailed	Cumulative	Double-tailed	Cumulative	Double-tailed
1	658	654	83	83	18	20
2	634	634	62	62	17	20
3	644	642	99	99	14	16
4	544	544	68	68	14	13
5	536	536	82	82	15	13
6	630	618	125	126	15	12
7	697	687	114	105	15	16
8	538	538	103	92	14	15
9	702	693	109	99	12	12
10	606	601	94	91	13	13
11	409	409	66	66	15	15
12	604	598	106	102	16	13
13	557	557	96	93	14	13
14	483	483	103	94	12	12
15	605	605	111	118	12	12
16	512	512	95	97	16	14
17	551	551	74	74	16	14
18	542	542	65	65	14	14
19	584	578	83	83	14	14
20	661	661	105	109	13	13

Table 8.9 - Uncertainty Values from Weighted Bayesian Method Using TxDOT Parking Lot Results

Test Pt.	Uncertainty as Standard Deviation (*ksi*)					
	AC		Base		Subgrade	
	Cumulative	Double-tailed	Cumulative	Double-tailed	Cumulative	Double-tailed
1	65	70	24	24	2	3
2	62	63	10	10	2	3
3	63	65	17	18	1	1
4	63	63	11	11	1	1
5	61	61	17	17	1	1
6	62	67	25	36	1	1
7	72	84	19	17	1	1
8	59	59	16	16	1	1
9	75	88	18	16	1	1
10	60	61	17	16	1	1
11	41	41	15	15	2	2
12	60	60	18	24	2	1
13	59	59	18	15	1	1
14	55	55	17	16	1	1
15	60	59	21	33	1	1
16	57	57	15	19	1	1
17	58	58	17	17	2	1
18	55	55	12	12	1	1
19	62	58	25	25	1	1
20	65	65	24	29	1	1

Table 8.10 - TxDOT Statistical Weighted Average Results

Test Pt.	Modulus (*ksi*)					
	AC		Base		SG	
	Avg	StDev	Avg	StDev	Avg	StDev
1	659	66	63	9	15	1.1
2	634	63	61	9	15	1.1
3	644	64	84	13	13	0.9
4	525	53	63	10	12	0.8
5	520	52	74	11	13	0.9
6	629	63	110	17	13	0.9
7	704	70	100	16	14	1.0
8	527	53	97	16	13	0.9
9	711	71	99	16	12	0.8
10	604	60	84	13	11	0.8
11	408	41	61	9	10	0.7
12	601	60	98	15	11	0.8
13	547	55	87	13	13	0.9
14	472	47	93	14	12	0.8
15	605	61	107	16	12	0.8
16	502	50	88	14	11	0.8
17	543	54	64	10	11	0.8
18	538	54	66	10	11	0.8
19	570	57	63	9	11	0.8
20	661	66	104	16	12	0.8

DATA FUSION OF TTI RIDE RUT SITE

The Ride Rut facility located at the Riverside Annex of Texas A&M University was another ideal site at this stage of the project. Figure 8.2 provides an idealized illustration of the pavement cross-section. The site consisted of 2 in. of ACP over a granular base, over an 8-in.-thick lime treated subbase. The subgrade consisted of highly plastic clay. The test section was 2000 ft long. The base thickness for the first and last 200 ft of the section was nominally 6 in. thick, and for the middle 1600 ft 14 in. thick. As reflected in Figure 8.2, twenty five points were tested at this site. The first five and the last four points were 50 ft apart to cover the 6-in.-thick base, and the middle sixteen points, which were located on the thicker base, were 100 ft apart.

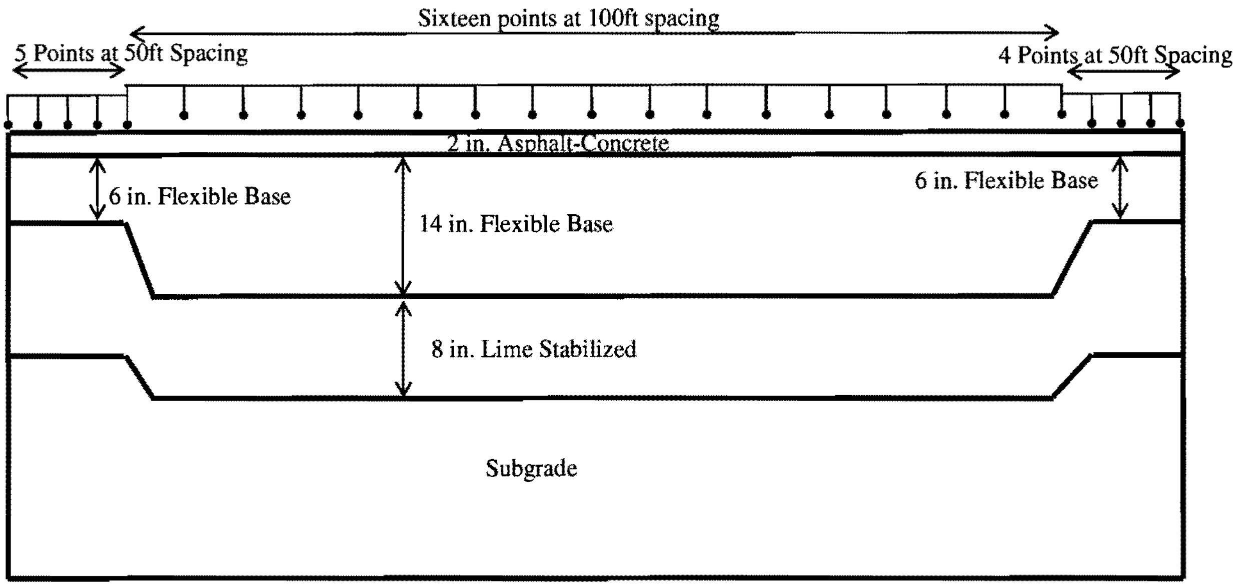


Figure 8.2 - Schematic of the test setup and test section at the Ride Rut facility

The GPR, FWD, SPA and PSPA were used at this site. This site provided us with the opportunity to test the feasibility of using JIM, as well as to implement the data fusion process. Each point was tested with the GPR first. After the GPR data was collected, the FWD followed by the SPA and PSPA were used at the site.

The procedure used to analyze the data can be summarized in following steps:

Step 1: Determine the seismic modulus profile from SPA using SMART software.
Step 2: Determine design modulus profile using the results of Step 1 in SMART software.
Step 3: Determine modulus profile from FWD using MODULUS.
Step 4: Determine the design modulus profile using the measured dispersion curves and the deflection basins using JIM.
Step 5: Extract discrete thickness for each test location from GPR data using COLORMAP
Step 6: Determine the modulus of ACP from PSPA data

The raw data that were fused from the TTI site are shown in Table 8.11. More detail about the analysis can be found in Abdallah et al. (2003) in the first report of this project.

Table 8.11 - TTI Data for Fusion

Test Pt.	Modulus (*ksi*)									
	AC				Base			Subgrade		
	M*†	S†	J†	P†	M	S	J	M	S	J
1	500	642	821	471	147	44	31	14	12	9
2	500	564	502	473	164	64	69	14	28	10
3	500	521	659	480	320	58	48	16	22	11
4	500	583	555	495	211	42	76	16	50	11
5	500	604	519	518	125	47	72	17	44	12
6	500	464	525	459	135	81	46	19	59	13
7	500	452	575	521	132	38	52	20	43	14
8	500	497	401	489	96	47	95	23	24	16
9	500	593	371	528	170	46	50	22	24	15
10	500	564	486	550	97	60	53	23	49	15
11	500	501	577	481	116	67	56	21	48	14
12	500	501	516	478	117	44	62	24	16	17
13	500	515	460	444	98	49	67	23	35	16
14	500	464	529	459	83	43	42	17	40	12
15	500	503	487	535	136	42	56	20	39	14
16	500	427	421	523	70	40	44	19	42	13
17	500	456	449	503	80	24	35	23	21	16
18	500	468	352	506	106	26	41	25	19	17
19	500	740	434	542	61	46	21	16	4	13
20	500	497	452	559	33	52	20	14	26	11
Avg	500	528	505	501	125	48	52	19	32	13
StDev	0	76	105	33	62	13	19	4	15	2
CoV	0%	14%	21%	7%	49%	28%	36%	19%	46%	17%

* Omitted from fusion † M=MODULUS, S = SMART, J = JIM, P = PSPA

Fuzzy Logic Fusion

Starting with the fuzzy logic method, the set boundaries must be defined. The same approach could almost be taken. The set boundaries shown in the lower half of Table 8.12 are centered about the rounded mean. Each set is one standard deviation wide and has the same shape pattern as was used for the TxDOT Parking Lot data, Z, Λ, Λ, Λ, S. In Table 8.12, the lower limit of the lowest set (for the base) is a negative number. Four options are available to deal with this potential problem. The first is to simply ignore this problem; since no data has a negative value, the lower set will be essentially void. The second option is to narrow the spread in the sets, that is reduce the standard deviation by a factor. The third option is to horizontally shift all of the sets to the right such that the lower limit of the lowest set is a non-negative value (it could still be set to zero). Finally, the sets can be defined manually in any manner desired. Similarly, the lowest set for the subgrade layer has its lower limit at zero and can be dealt with in a similar fashion.

Table 8.12 - Initial Set Boundaries for TTI Data

Parameter	**Modulus (*ksi*)**		
	AC	**Base**	**SG**
Average	500	75	20
StDev	75	45	10
Set	Membership Boundaries (*ksi*)		
Very Low	350	-15	0
Low	425	30	10
Average	500	75	20
High	575	120	30
Very High	650	165	40

To deal with the negative and zero values of the initial guess for the sets, the width of the base (i.e. the standard deviation) is changed to 35 *ksi* and for the subgrade ranges to 5 *ksi*. The revised set boundaries are shown in Table 8.13.

Table 8.13 - Revised Set Boundaries for TTI Data

Parameter	**Modulus (*ksi*)**		
	AC	**Base**	**SG**
Average	500	75	20
StDev	75	35	10
Set	Membership Boundaries (*ksi*)		
Very Low	350	5	10
Low	425	40	15
Average	500	75	20
High	575	110	25
Very High	650	145	30

The weights for each layer are defined in Table 8.14. For this case, below average AC moduli are weighted less than the above average moduli. Some very low moduli are estimated for the base and the subgrade layers. As these values appear to be outliers, the weight applied to the lowest sets for these layers are reduced so that the fused outcome is not unfairly biased toward these outliers.

Table 8.14 - Weights for Membership Functions Used for TTI Site

Layer	Very Low	Low	Average	High	Very High
AC	1.5	2	3	2.5	2
Base	0.5	2	3	2.5	1.5
Subgrade	0.5	2	2.5	2.25	2

The fuzzy sets for the TTI site are shown in Figure 8.3. The sets are ordered from left to right. Additional device modification values are assigned to the PSPA as indicated in Table 8.15. Since PSPA does not yield moduli for the base and subgrade layers, no modification values are assigned to them.

Table 8.15 - Device Modification Values for TTI Site Data

Method	Device Modification Factor		
	AC	Base	Subgrade
MODULUS	0	1	1
SMART	1	1	1
JIM	1	1	1
PSPA	1	--	--

The results of applying the fuzzy logic method are shown in Table 8.16 using the raw data from Table 8.11 and the additional input information from Tables 8.13 through 8.15.

Hybrid Fusion

In the hybrid method, the same set boundaries defined for the TTI fuzzy logic were used as shown in Table 8.13. The modification values for the coefficients of variation will be defined as shown in Table 8.17. The set weights function as the modification factors for the coefficients of variation. The set weights were defined towards weighting the low data for the AC modulus fairly high and then reducing the weight of the low values for the base and subgrade layers.

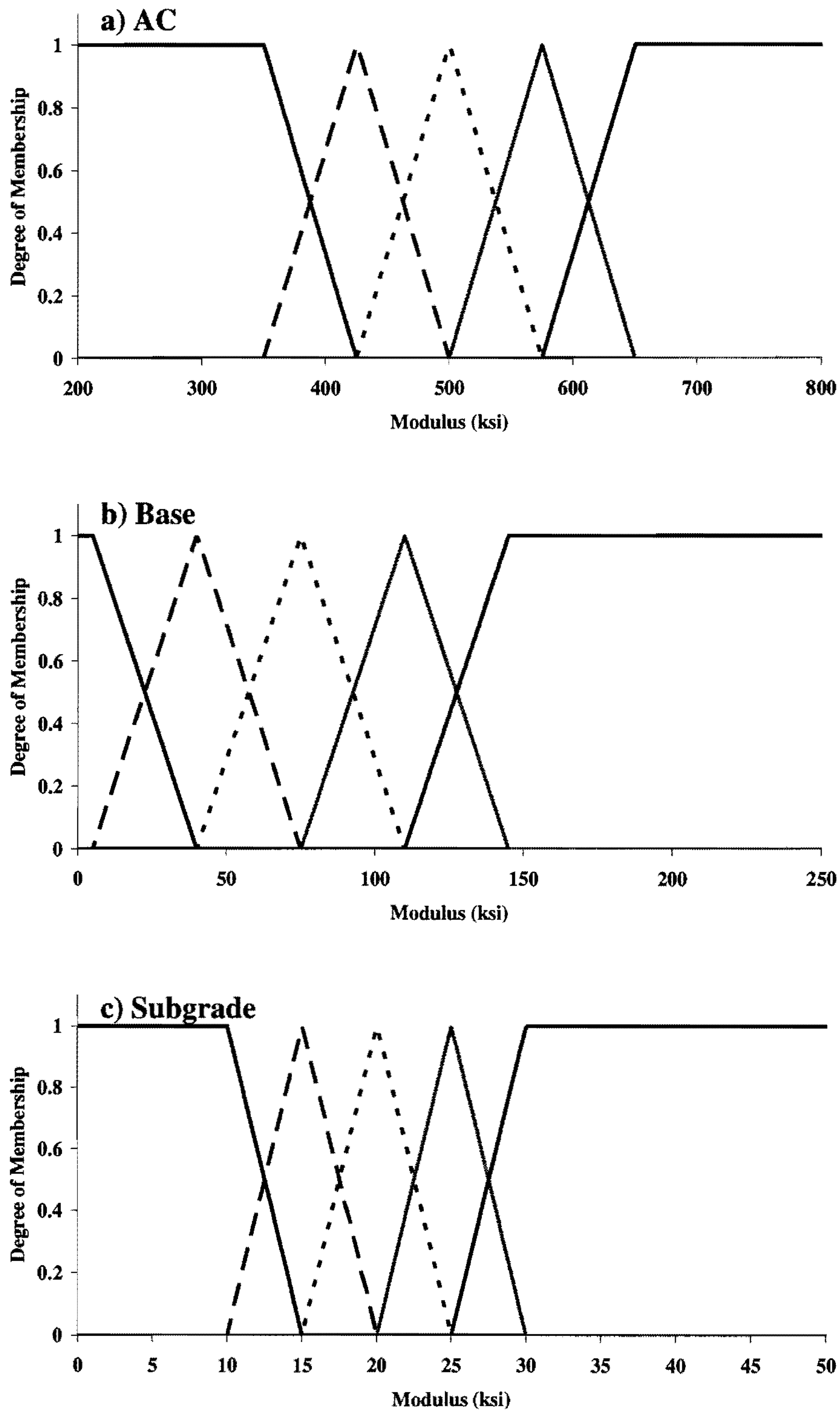


Figure 8.3 – Fuzzy Sets for TTI Site

Table 8.16 - Results for TTI Data from Fuzzy Logic

Test Pt.	Modulus (*ksi*)								
	AC			Base			Subgrade		
	Avg	StDev	CoV	Avg	StDev	CoV	Avg	StDev	CoV
1	633	61	10%	70	22	31%	12	2	19%
2	513	46	9%	87	21	24%	20	6	32%
3	546	49	9%	117	39	33%	18	5	26%
4	543	48	9%	96	27	28%	29	9	32%
5	545	48	9%	80	21	26%	26	8	29%
6	484	45	9%	84	20	24%	31	9	30%
7	517	46	9%	72	21	30%	26	7	25%
8	469	37	8%	81	21	26%	21	4	19%
9	512	39	8%	79	23	29%	20	4	20%
10	533	45	8%	71	19	26%	28	7	25%
11	519	47	9%	78	20	25%	28	7	26%
12	499	45	9%	75	21	28%	19	3	16%
13	475	42	9%	72	19	27%	24	5	22%
14	485	45	9%	59	18	30%	25	7	28%
15	508	43	9%	74	21	28%	25	6	24%
16	463	35	8%	54	15	29%	26	7	27%
17	471	38	8%	55	20	36%	20	4	19%
18	456	33	7%	65	23	35%	20	4	17%
19	567	45	8%	48	14	29%	13	2	13%
20	504	41	8%	39	9	22%	19	6	30%
Avg	512	--	--	73	--	--	23	--	--
StDev	42	--	--	17	--	--	5.0	--	--
CoV	8%	--	--	24%	--	--	22%	--	--

Table 8.17 - TTI Modification Values for Coefficients of Variation

Layer	Fuzzy Set Weights				
	Very Low	Low	Average	High	Very High
AC	0.85	0.95	1	0.9	0.8
Base	0.5	0.85	1	0.9	0.7
Subgrade	0.5	0.85	1	0.8	0.7

The results from applying the hybrid method to the TTI data are shown in Table 8.18. The hybrid fused values were more conservative than the fuzzy logic fused values.

Table 8.18 - Results for TTI Data from Hybrid Method

Test Pt.	Modulus (*ksi*)								
	AC			Base			Subgrade		
	Avg	StDev	CoV	Avg	StDev	CoV	Avg	StDev	CoV
1	490	30	6%	39	10	26%	11	2	20%
2	485	29	6%	67	15	22%	12	2	20%
3	491	29	6%	54	13	24%	14	2	17%
4	507	30	6%	50	12	24%	14	2	17%
5	526	32	6%	55	13	23%	15	2	16%
6	464	27	6%	62	14	23%	16	2	15%
7	510	30	6%	43	11	25%	17	2	14%
8	482	28	6%	57	13	23%	18	3	14%
9	515	32	6%	49	12	24%	17	2	14%
10	545	34	6%	60	13	22%	17	3	15%
11	488	29	6%	64	14	22%	17	2	15%
12	484	28	6%	51	12	23%	19	3	14%
13	454	27	6%	57	13	22%	18	3	14%
14	464	27	6%	45	10	23%	15	2	16%
15	525	31	6%	48	11	24%	17	2	14%
16	493	29	6%	44	10	23%	16	2	15%
17	490	28	6%	29	8	28%	18	3	14%
18	483	28	6%	32	9	27%	19	3	14%
19	541	34	6%	33	9	26%	11	2	17%
20	534	33	6%	31	8	27%	13	2	19%
Avg	498	--	--	49	--	--	16	--	--
StDev	26	--	--	11	--	--	2	--	--
CoV	5%	--	--	24%	--	--	16%	--	--

Bayesian Fusion

The input for the Bayesian methods was simply the data shown in Table 8.11. The results of applying the Bayesian methods are shown in Tables 8.19 through 8.21.

Table 8.19 - Modulus Values from Decision Bayesian Method Using TTI Site Results

Test Pt.	Modulus (*ksi*)					
	AC		Base		Subgrade	
	Cumulative	Double-tailed	Cumulative	Double-tailed	Cumulative	Double-tailed
1	642	471	44	44	13.5	13.5
2	564	502	64	69	28	28
3	480	480	58	48	16.4	16.4
4	495	495	42	42	16.2	16.2
5	518	518	47	47	12	12
6	525	525	46	46	13	13
7	521	521	52	52	14	14
8	489	489	47	47	16	16
9	528	528	46	46	15	15
10	550	486	60	53	15	15
11	481	481	56	56	14	14
12	478	478	62	44	17	23.6
13	515	515	49	49	16	16
14	529	529	43	43	12	12
15	535	503	56	42	14	14
16	523	523	40	40	13	13
17	503	503	35	35	16	16
18	506	506	41	41	17	17
19	542	434	46	46	13	13
20	559	497	52	52	11	11

Table 8.20 - Modulus Values from Weighted Bayesian Method Using TTI Site Results

Test Pt.	Modulus (*ksi*)					
	AC		Base		Subgrade	
	Cumulative	Double-tailed	Cumulative	Double-tailed	Cumulative	Double-tailed
1	563	483	67	58	12	12
2	524	510	80	91	22	22
3	530	502	55	52	15	15
4	534	516	76	48	24	15
5	537	526	65	62	20	16
6	501	496	75	70	17	16
7	530	522	63	61	19	17
8	486	486	56	53	19	20
9	544	535	65	54	18	18
10	543	526	60	60	20	18
11	518	501	67	70	19	17
12	497	496	61	60	20	20
13	494	494	59	57	20	24
14	504	498	46	46	19	17
15	521	505	63	61	19	18
16	513	498	44	44	19	17
17	493	492	48	48	19	19
18	500	498	60	60	20	20
19	532	481	46	46	13	13
20	517	482	51	51	16	16

Table 8.21 - Uncertainty Values from Weighted Bayesian Method

Test Pt.	Modulus (*ksi*)					
	AC		Base		Subgrade	
	Cumulative	Double-tailed	Cumulative	Double-tailed	Cumulative	Double-tailed
1	50	29	19	14	2	2
2	49	47	17	24	8	8
3	41	39	12	12	2	2
4	39	31	18	11	6	2
5	34	40	14	14	4	2
6	70	63	22	21	2	2
7	37	34	16	15	3	2
8	29	29	12	13	2	3
9	35	33	15	11	2	2
10	39	61	12	14	3	2
11	46	34	14	19	3	2
12	43	41	13	13	3	3
13	57	57	12	12	3	6
14	71	62	9	9	4	2
15	32	48	15	13	3	2
16	32	31	9	9	3	2
17	29	29	14	14	2	3
18	30	30	19	19	2	4
19	33	54	12	12	2	2
20	35	55	14	14	4	4

Statistical Weighted Average Fusion

Finally, the statistical weighted average method was applied to the TTI data. The results of this fusion are shown in Table 8.22.

Table 8.22 - TTI Statistical Weighted Average Results

Test Pt.	Modulus (*ksi*)					
	AC		Base		SG	
	Avg	StDev	Avg	StDev	Avg	StDev
1	498	29	38	8	10	1.3
2	486	28	69	14	11	1.4
3	494	28	55	12	13	1.7
4	510	29	49	11	13	1.7
5	529	30	55	12	14	1.8
6	464	27	61	13	15	2.0
7	510	29	43	9	16	2.1
8	481	28	56	12	18	2.3
9	513	30	49	10	17	2.2
10	546	31	60	12	17	2.2
11	489	28	65	13	16	2.1
12	484	28	50	11	19	2.3
13	454	26	56	11	18	2.3
14	464	27	45	9	14	1.8
15	525	30	47	10	16	2.1
16	492	28	43	9	15	2.0
17	490	28	28	6	18	2.3
18	478	28	30	6	19	2.3
19	546	32	29	6	9	1.3
20	536	31	27	6	12	1.6

DISCUSSION OF RESULTS

TxDOT Parking Lot Results

The examination of the results as presented in this chapter is performed both in tabular and graphical form. Table 8.23 shows the representative averages and uncertainties in moduli for the TxDOT Parking Lot from each method.

Table 8.23 - Summary of TxDOT Site Statistics from Fusion Methods

Method	Modulus (*ksi*)					
	AC		Base		SG	
	Avg	StDev	Avg	StDev	Avg	StDev
Statistical Weighted Average	580	79	83	17	12	1
Fuzzy	580	73	99	27	14	1
Hybrid	581	78	83	17	12	1
Cumulative Decision	588	67	87	13	14	5
Double-tailed Decision	584	66	88	17	14	4
Cumulative Fusion	585	74	92	18	14	2
Double-tailed Fusion	582	71	90	18	14	2
Average	554	52	98	32	15	1
MODULUS	500	0	89	41	18	2
SMART	583	82	122	51	18	3
JIM	580	80	83	15	10	1

Figures 8.4 through 8.6 show the variations of the moduli for each of the layers. In each figure, the top graph contains the point-by-point comparison of the moduli, the middle graph the point-by-point comparison of the uncertainty values, and the bottom graph the spread in the site

averages. For the site averages, no fusion method was used other than a regular average and standard deviation computation. The point-by-point values that are plotted are included in Appendix A.

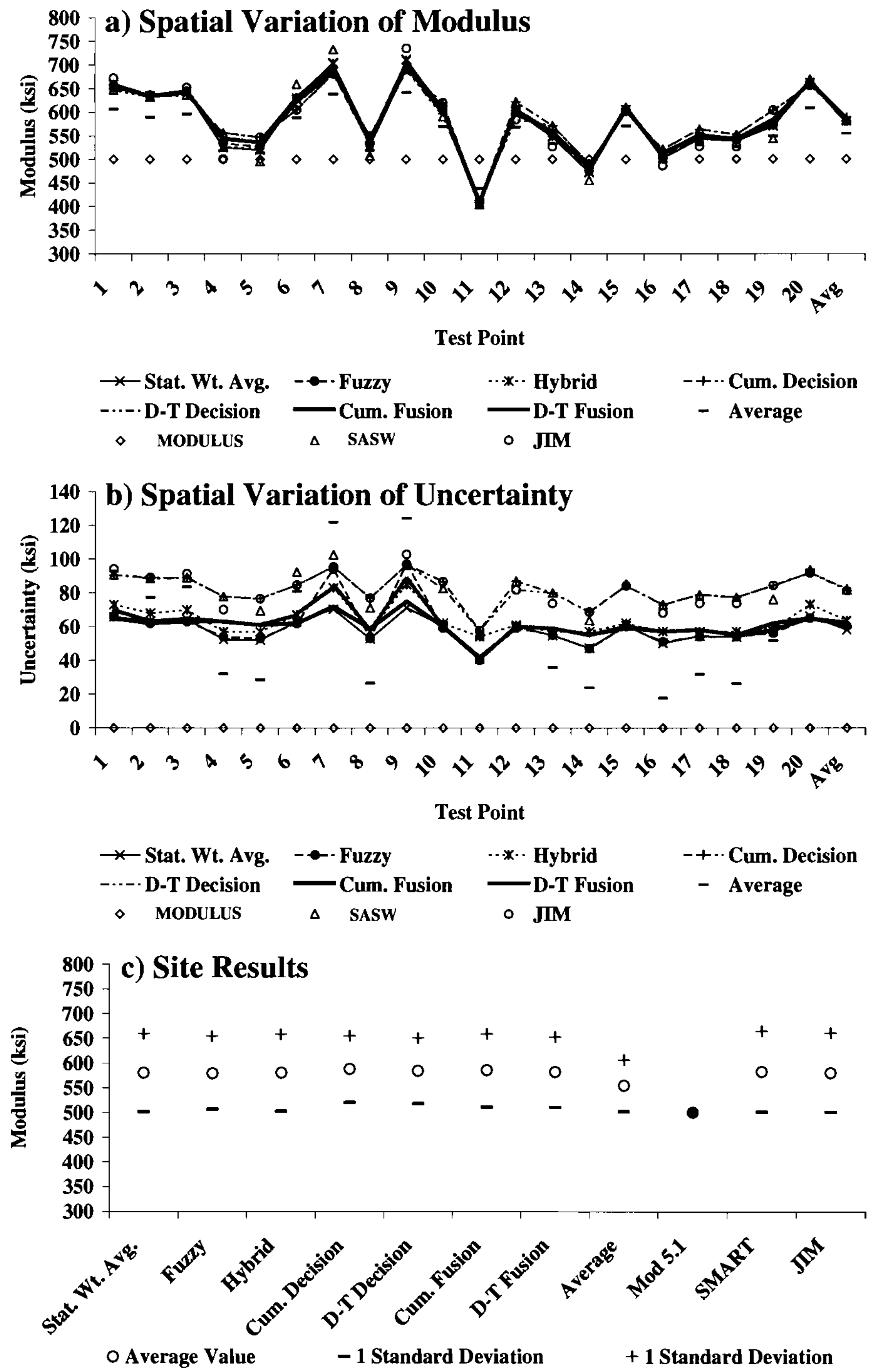


Figure 8.4 - Summary of Results for AC Modulus Using TxDOT Parking Lot Data

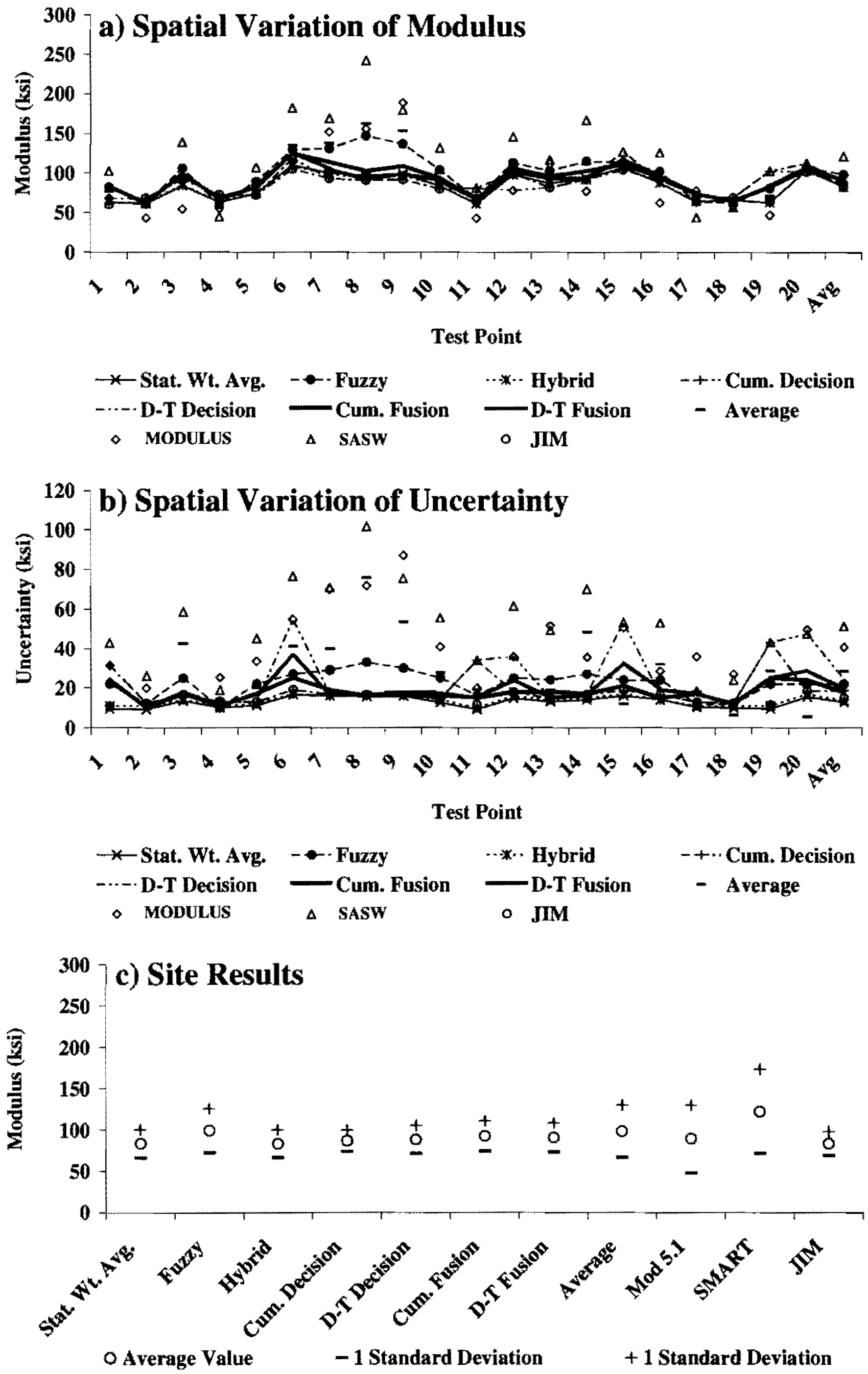


Figure 8.5 - Summary of Results for Base Modulus Using TxDOT Parking Lot Data

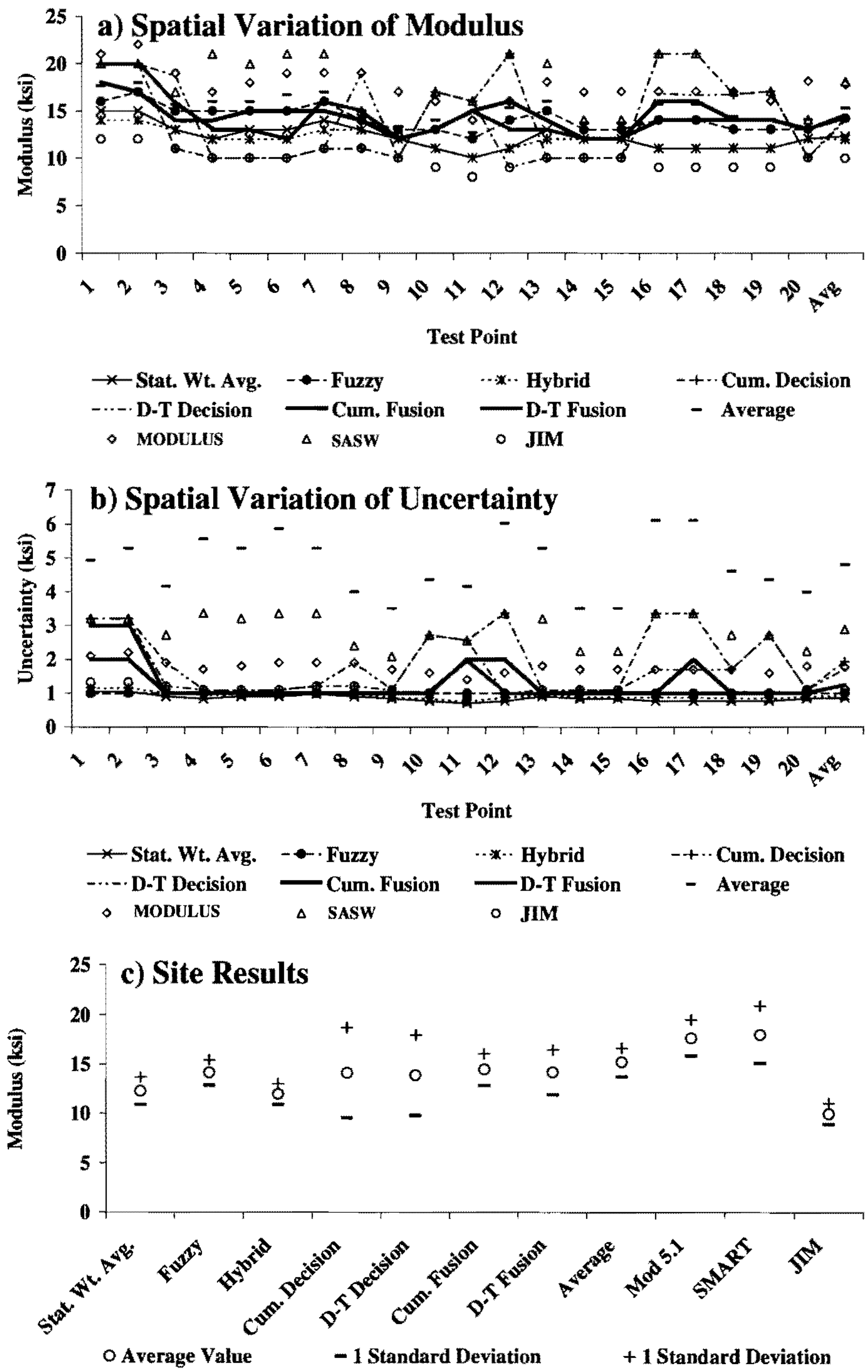


Figure 8.6 - Summary of Results for Subgrade Modulus Using TxDOT Parking Lot Data

From Figures 8.4 through 8.6, a comparison of the fusion methods can be made. Each layer will be discussed separately. From Figure 8.4a, the spatial variation in the AC modulus can be assessed. The moduli obtained by the non-decision methods (Bayesian decision) lie somewhere between the maximum and minimum moduli measured at each test point. Some smoothing of the moduli along the site is evident. The final output (fused value) based on any of the methods is obtained by combining data from more than one source, hence, a decision is not blindly taken based solely on the preference of a method or a simple average. One key observation is that moduli that are extreme (very high or very low) are closer to the average when combined with data from other sources. Any fused value will lie between the maximum and minimum of the values of the points used to generate the fused result.

Figure 8.4b shows the uncertainty associated with each method at each point. The uncertainty for the moduli from MODULUS is numerically zero since it was fixed during the data reduction process. The fusion methods have the lowest uncertainty for about half the points tested. For the other points, it appears that the simple average method yields the lowest uncertainty, however, this may be because the simple average method included the fixed modulus "estimated" by MODULUS. The simple standard deviation used for the simple average method would not include any information about what the expected site error may be (via site statistics) as the fusion methods did. The uncertainty in that average method is based entirely on the difference of the three data points. Also, consider that taking a standard deviation of an average of only three data points is not the most reliable way to find the uncertainty for such a small set of numbers. Of the fusion methods, it appears that the fuzzy and statistical weighted average methods seem to have the lowest uncertainty in the majority of the twenty test points considered.

Figure 8.4c provides information on the site statistics based on the different methods, including the raw (unfused) results. The standard deviations for each method for the site do not differ significantly ranging from only 9% to 14% (refer to Appendix A, Table A.1). Additionally, the final site results are all within the ballpark of one another, with the simple average being lower. Inspecting the three graphs as a whole, the overall site average may not significantly change. The uncertainty associated with individual moduli at points can be defined with more confidence after a numeric process is performed to obtain a more informed result from more than one source.

In Figure 8.5, some conclusions about the base layer can now be made. From Figure 8.5a, the erratic behavior of the raw moduli is reduced considerably. Much of this was due to the influence of the JIM on the fusion process. The JIM results were less erratic as compared to either MODULUS or SMART results and as a whole, tended to be more conservative as expected. In this case, the representative moduli form the fuzzy method was influenced more by the higher values than the other methods, yet overall, the fuzzy results are at or near those of the other fusion methods.

Figure 8.5b provides an assessment of the reliability of the fused and unfused data at each point. The fusion methods have a considerable advantage when the uncertainty in the result is looked at. The non-fused data points have uncertainties as much as five times higher than some of the fused data points. The reliability of the JIM is highly influential on the composite uncertainties from the fusion methods. Even though the JIM uncertainty is lower than some of the fusion methods, namely the fuzzy method, the results of the fuzzy method include data from more than one source. Accepting the value with the lowest uncertainty may be an option, but making that decision may result in useful or good data being omitted or eliminated without just cause.

The representative moduli from all fusion methods for the base layer are compared in Figure 8.5c. With the exception of the joint inversion method, all of the fusion methods have a lower coefficient of variation than any of the raw data methods or the simple average.

The moduli for the subgrade layer are shown in Figure 8.6. Figure 8.6a shows the spatial variation in the modulus profiles for the site. With few exceptions, the fusion methods mirror the JIM results, which in turn mirror the MODULUS. This is due to the nature of the JIM as described by Abdallah *et al.*(2003). The spread in the raw data is not too high in the raw data either, ranging from 10% to 16% (Table A.3). As indicated before, FWD provides information of the subgrade substantially deeper than the SMART method.

The point-by-point uncertainty of different methods shown in Figure 8.6b can aid in the comparison of the methods. Overall, the uncertainty of the fusion methods is much lower than those of the unfused methods when observed point-by-point. The (Bayesian) decision methods tend to exhibit higher uncertainties, as they reflect the uncertainty of the raw data method decided on at that point.

The site results reflected in Figure 8.6c show that the fused results have a lower uncertainty than the raw data with the highest uncertainty (SMART) even though the SMART data was included in the fused result. Based on the coefficient of variation, only the decision methods have values higher than 20%, with values up to 33% (Table A.3). The coefficients of variation for the other fusion methods are all less than the 16% of the SMART method, and in most cases are very near or even below the lowest uncertainties of the input data.

TTI Site Results

Table 8.24 shows the site statistics for the fusion of the TTI data. PSPA data is only possible for the AC layer and thus no values are shown for either of the base or subgrade layers. Figures 8.7 through 8.9 show the raw and fused data from the TTI site.

Figure 8.7a shows how the fusion methods can smooth the point by point results. At this site, even the decision methods smoothed out the modulus profiles. Values than seemed to be much higher or lower than the typical values for a device over the site were not selected with either of the decision methods.

Table 8.24 - Summary of TTI Site Statistics from Fusion Methods

Method	Modulus (*ksi*)					
	AC		Base		SG	
	Avg	StDev	Avg	StDev	Avg	StDev
Statistical Weighted Average	499	27	48	12	15	3
Fuzzy	512	42	73	17	23	5
Hybrid	498	26	49	11	16	2
Cumulative Decision	524	38	49	8	15	4
Double-tailed Decision	499	24	47	7	15	4
Cumulative Fusion	519	20	60	10	19	3
Double-tailed Fusion	502	16	58	11	18	3
Average	508	35	75	25	22	5
MODULUS	500	0	125	62	19	4
SMART	528	76	48	13	32	15
JIM	505	105	52	19	13	2
PSPA	501	33	--	--	--	--

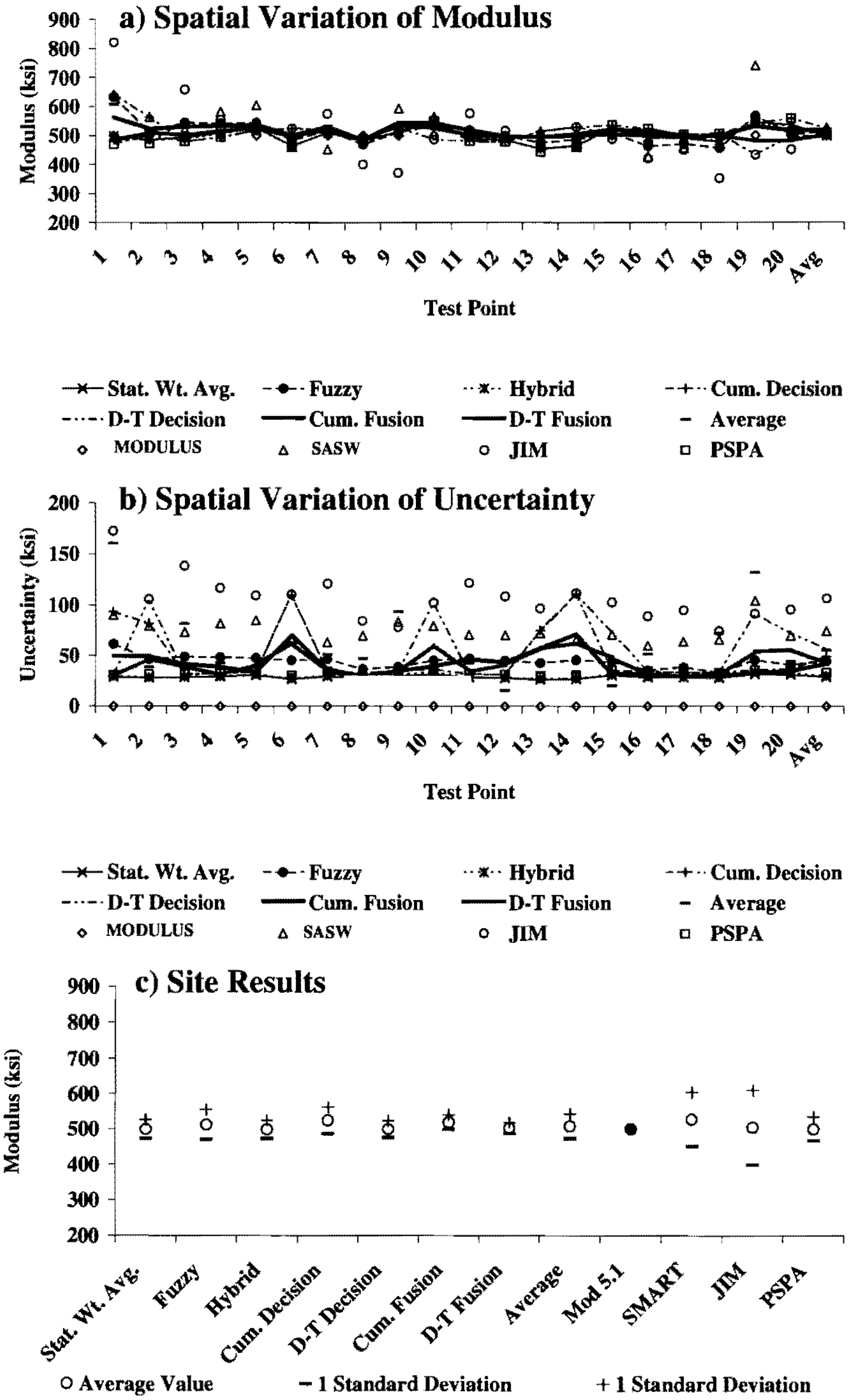


Figure 8.7 - Summary of Results for AC Modulus Using TTI Site Data

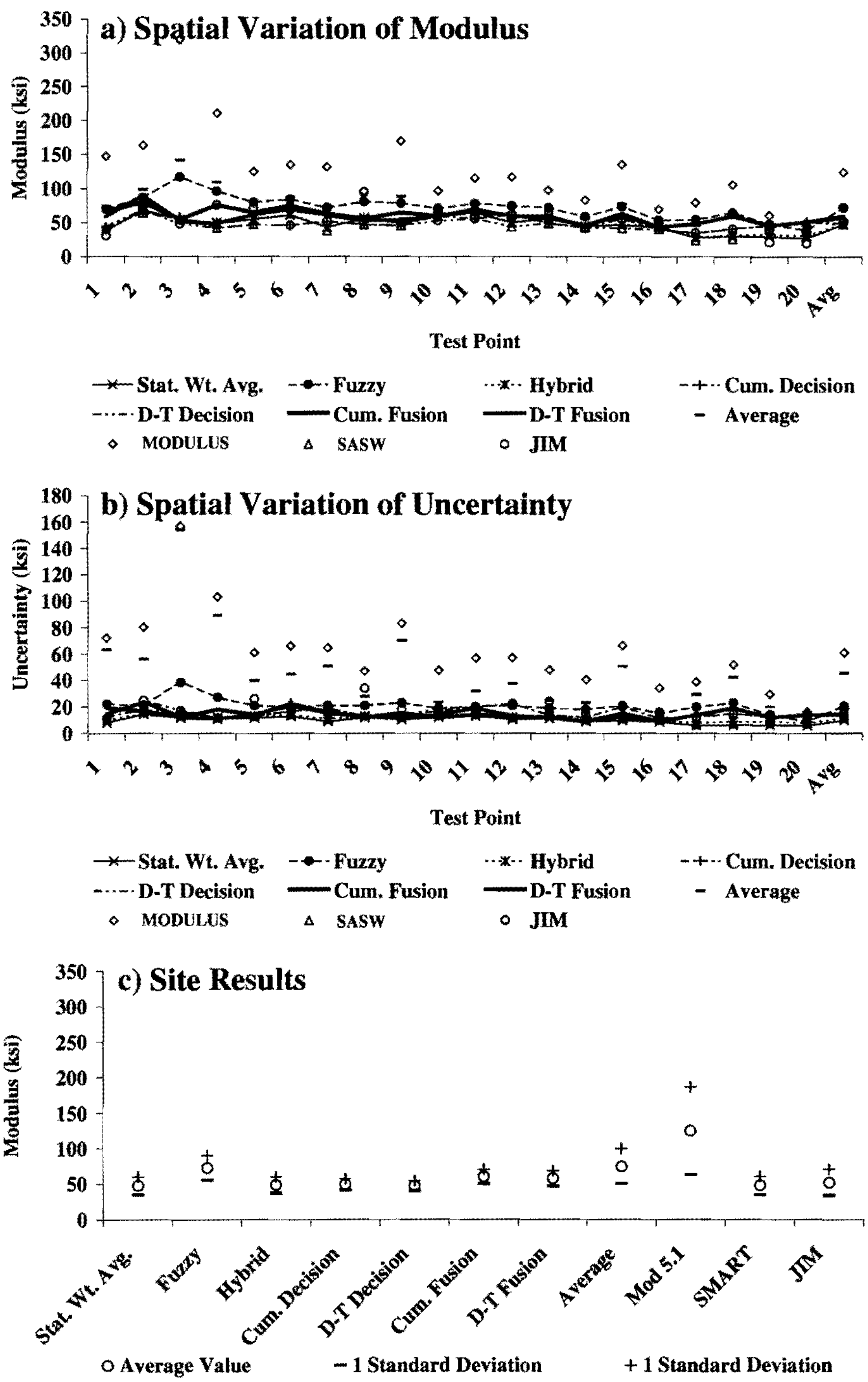


Figure 8.8 - Summary of Results for Base Modulus Using TTI Site Data

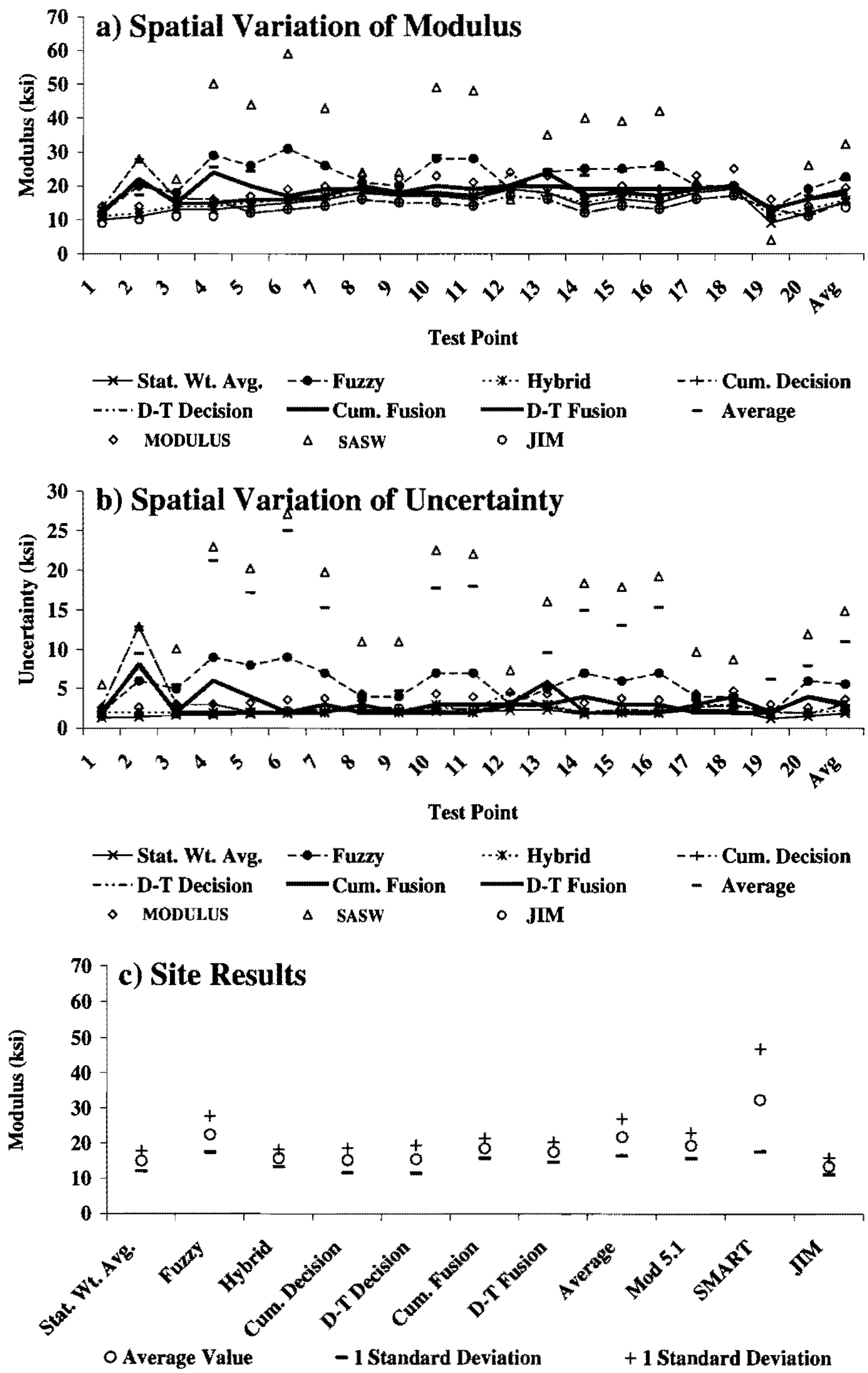


Figure 8.9 - Summary of Results for Subgrade Modulus Using TTI Site Data

The uncertainty values in Figure 8.7b are the lowest for the fusion methods and closely shadow the uncertainty of the PSPA which had the lowest uncertainty of any of the methods. Some of the uncertainty values are quite high, particularly with the simple average method at Points 1 and 19. The decision methods seem to have the next highest uncertainty, and the fusion methods all appear to be very close in uncertainty, with the occasional spike (Points 7 and 14). The spikes occurred where the Bayes cumulative fusion was employed; however, the uncertainties for that method at all other points are quite stable and comparable to the other fusion methods.

Overall, the site results (site average of point fusion) are quite similar and all have relatively low site standard deviations with the exception of the decision methods, where the site uncertainties are about twice that of the fusion methods as is evident from Figure 8.7c. Still, the decision uncertainty is a median representation of the uncertainty if it is compared to the raw data. The values are not quite as high or low as the highest and lowest uncertainties of the raw data.

The fusion of the base data that is plotted in Figure 8.8a shows that the data was greatly smoothed out by the fusion methods. On the other hand, it shows that one of the raw methods greatly controlling the fusion process. In this case, the SMART and JIM methods are very close and thus supported each other significantly. In that process, the fused results are closer to the midpoint between the two methods. With the exception of about four points (3, 4, 9, and 15), the trend of the MODULUS outputs is very similar to the trend of the SMART and JIM data, only shifted up from about 50 to 100 *ksi*. While the raw data agreed on the trend of the data in most cases, there was some conflict with the magnitude. In cases where a high-end estimate resulted while some of the data was considerably lower, some considerations can be made in the assigning of weights to low moduli as is allowed in the fuzzy and hybrid methods and to some extend in the statistical weighted average method.

The uncertainty of the base layer, shown in Figure 8.8b, exhibits a trend similar to the moduli in Figure 8.8a. High uncertainties are evident in the simple average and MODULUS results at Test Points 3, 4, 9, and 15. Taking a simple average equally weighs the moduli from each method, including the erratic data from MODULUS. The fusion methods proved to significantly reduce the uncertainty for the base layer.

The plots in Figure 8.7c simply reiterate what has already been shown by Figures 8.8a and 8.8b. The fusion methods all yielded consistent results for the base modulus, which were furthermore close in value to each other. The site averages reflect that consistency (as per stability) and closeness (as per magnitude).

The spatial variations in modulus of subgrade are shown in Figure 8.9a. The fused profile is fairly stable. As was evident with the base, the data was controlled by one or more of the input methods. The methods that controlled the subgrade were MODULUS and JIM. SMART method was more variable. It seems that the fuzzy method is quite affected by the SMART results. Changes to the device modification values and/or the membership functions of the fuzzy sets could have been used to obtain a more conservative estimate instead of weighting the moduli from SMART as much as they were weighted. Overall though, the fuzzy results are more stable than the SMART values. The other fusion methods are grouped much closer to one another and to MODULUS and JIM methods, which were the controlling methods for all of the fusion methods.

An examination of the uncertainties in Figure 8.9b supports some of the observations made about the SMART method and its impact on the fuzzy results. The uncertainty for the fuzzy method was much higher relative to the other fusion methods than observed in the overlaying layers and the results from the other site. As can be expected, the simple average method yielded high uncertainty as it mirrored the SMART results. The remainder of the fusion methods yielded uncertainties near the MODULUS and JIM uncertainties aside from a handful of points.

In comparing the site averages and standard deviations using Figure 8.9c, the conclusions that were reached for the profile carry over to the trend of the averages. The average SMART modulus was high for a subgrade layer. The data cannot be disregarded completely since these results are more indicative of the semi-stabilized subgrade. The MODULUS and JIM results had fairly tight ranges, as did most of the fusion methods. The decision methods had a slightly higher spread than all of the non-decision methods except for the Fuzzy method. The site average for the Fuzzy method is still reasonable for a subgrade and the data included from the SMART component of that fusion may have validity, as the SMART process in a way was influenced by the properties of the semi-stabilized layer that was overlooked by the other methods. However, the other fusion methods were not affected so heavily by the SMART data.

CHAPTER NINE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This report attempted to show the feasibility of using data fusion techniques in the analysis of nondestructive test data for pavements. This feasibility not only includes the methods for the data fusion, but also the considerations that must be taken if data fusion is to be applied to NDT of pavements.

An overview of data fusion methods are introduced in Chapter 3. Simple examples are used to present and illustrate fusion algorithms. In Chapters 4 through 7, the process of implementing each data fusion method for analyzing nondestructive pavement testing data is offered. Chapter 4 presents an algorithm of applying data fusion to NDT results using statistical weighted average method. The adoption of statistical weighted average method to NDT data is the easiest to incorporate since only simple statistics are required to develop the algorithm. In Chapter 5, the fuzzy logic approach to fusing NDT data is developed. Fuzzy logic is perhaps the most complex fusion method. However, parameters regarding the inherent error in NDT analysis methods and uncertainty and variability associated in NDT data were incorporated in the algorithm. The algorithm was developed with two levels of user involvement: a) manual and automatic. The manual process allows users to interface with several aspects of setting up the fussy logic set and membership functions. For novice users, the automatic algorithms provides typical shapes and functions for users to select from and which aides in fusing NDT data with little knowledge in fuzzy logic. Chapter 6 provides detail of the development of the hybrid method and its applicability to fuse NDT data. The main approach used to developing the hybrid algorithm is by combining the fuzzy logic method with the weighted average method. The last of the data fusion methods that is explored is the Bayesian inference method presented in Chapter 7. This method requires prior probability and conditional densities. For this reason applying this method to fusing NDT data was not as uncomplicated as the other methods, and therefore was abandoned for this project. Chapter 8 provides a presentation of results of fusing the results of popular NDT analyses for two sites. The results using all the fusion algorithms presented in this report are presented. It is important to point out that the use of the fuzzy logic method is most appropriate considering its flexibility of combining both statistical and expert inputs into the fusion process. This method also seems practical for applying to NDT data.

The methods presented in this study show that there are many options for the potential fusion of data from multiple devices that are used to nondestructively test pavements. A protocol for verifying the results of the fusion processes would be required to see which method is the most appropriate.

REFERENCES

- Abdallah, I., Yuan, D., Nazarian, S., and Ferregut C. (2003). "Feasibility of Integration of Non-Destructive Testing Data Analysis Techniques" *Center for Transportation Infrastructure Systems: Research Report No. 0-4393-1*, The University of Texas at El Paso, El Paso, Texas.

- Abdallah, I., Yuan, D., and Nazarian, S. (2002). "Validation of Software Developed for Determining Design Modulus from Seismic Testing." *Center for Highway Materials Research: Research Report No. 0-1780-5*, The University of Texas at El Paso, El Paso, Texas.

- Ang, A. H-S. and Tang, W. H. (1975). *Probability Concepts in Engineering Planning and Design: Volume I – Basic Principles*, John Wiley and Sons, New York, New York.

- National Institute of Standards and Technology (NIST). (2001). http://www.itl.nist.gov/div898/software/dataplot/refman2/ch2/homepage.htm, November 2003.

- Favata, M. (2001). "PLCs and Fuzzy Logic." http://www.industrialtext.com/eBooks/Fuzzy_Logic.PDF. July 2003.

- Ganji, M. (2004). "Optimization of Nondestructive Test Results Using the Joint Inversion Method." Master's Thesis, Department of Civil Engineering, The University of Texas at El Paso, El Paso, Texas.

- Gros, X. E. (1997). *NDT Data Fusion*, Jon Wiley and Sons, London, England.

- Jantzen, J. (1997). "A Robustness Study of Fuzzy Control Rules." *Proceedings of Fifth European Congress on Fuzzy and Intelligent Technologies*, ELITE Foundation, Aachen, Germany, 1222-1227.

- Jantzen, J. (1998). "Tutorial on Fuzzy Logic." Department of Automation, Technical University of Denmark. http://fuzzy.iau.dtu.dk/download/logic.pdf. July 2003.

- Krzysztofowicz, R. and Long, D. (1990). "Fusion of Detection Probabilities and Comparison of Multisensor Systems." *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 20, No. 3, May/June 1990, 665-677.

- Development of a Common Educational and Training Infrastructure (DeCETI). "Multi-Source Information Fusion for Satellite Images Classification." http://www.survey.ntua.gr/main/labs/rsens/DeCETI/IRIT/MSI-FUSION/ June 2003.

- Luo, R. C. and Kay M. K. (1989). "Multisensor Integration and Fusion in Intelligent Systems." *IEEE Transactions of Systems, Man, and Cybernetics*, Vol. 19, No. 5, September/October 1989, 901-931.

- Maser, K. R., Scullion, T., and Briggs, R. C. (1991). "Use of Radar Technology for Pavement Layer Evaluation." *Texas Transportation Institute: Research Report No. 903-5F*, Texas A&M University, College Station, Texas.

- Maser, K. R. and Scullion, T. (1992). "Influence of Asphalt Layering and Surface Treatments of Asphalt and Base Layer Thickness Computations Using Radar." *Texas Transportation Institute: Research Report No. 1923-1*, Texas A&M University, College Station, Texas.

- "Fuzzy Logic Toolbox Tutorial" (Mathworks) (2003). http://www.mathworks.com/access/helpdesk/help/toolbox/fuzzy/fuzzytut.shtml. July 2003.

- Nazarian, S. and Desai, M. R. (1993), "Automated Surface Wave Method: Field Testing," Journal of Geotechnical Engineering Division, ASCE, Vol. 119, No. GT7, New York, NY, pp. 1094-1111.

- Nazarian, S., Yuan, D., and Baker, M. R. (1995). "Rapid Determination of Pavement Moduli with Spectral Analysis of Surface Waves Method." *Center for Geotechnical and Highway Materials Research: Research Report No. 1243-1*, The University of Texas at El Paso, El Paso, Texas.

- Nazarian, S., Baker, M. R., and Crain, K. (1993). "Development and Testing of a Seismic Pavement Analyzer." *Strategic Highway Research Program: Report No. SHRP-H-375*, National Research Council, Washington, D.C.

- Osegueda, R. A., Lopez, H., Pereyra, L., and Ferregut, C. M. (2000). "Localization of Damage Using Fusion of Modal Strain Energy Differences." Fast Center for Structural Integrity of Aerospace Systems, The University of Texas at El Paso, El Paso, Texas.

- Park, S. and Lee, C. S. G. (1993). "Fusion-based Sensor Fault Detection." *Proceedings of the 1993 Symposium on Intelligent Control*, Chicago, Illinois, 156-161.

- Ricci, E. A., Meyer, A., H., Hudson, W. R., and Stokoe, K. H. II. (1985). "The Falling Weight Deflectometer for Nondestructive Evaluation of Rigid Pavements." *Center for Transportation Research: Research Report No. 387-3F*, The University of Texas at Austin, Austin, Texas.

- Scullion, T., Lau, C-L., and Chen, Y. (1994). "Implementation of the Texas Ground Penetrating Radar System." *Texas Transportation Institute: Research Report No. 1223-1*, Texas A&M University, College Station, Texas.

- Starr, A., Hannah, P., Esteban, J., and Willetts, R. (2000). "Adapting Data Fusion Frameworks for Condition Based Maintenance." *Symposium on Condition Based Maintenance for Highly Engineered Systems*, Pisa, Italy.

- Stokoe, K. H., II, Hudson, W. R., and Miner, B. F. (1991). "The Falling Weight Deflectometer and Spectral Analysis of Surface Waves Test for Characterizing Pavement Moduli." *Center for Transportation Research: Research Report No. 1123-7F*, The University of Texas at Austin, Austin, Texas.

- Tanaka, K. (1991). *An Introduction to Fuzzy Logic for Practical Applications*, Springer, New York, New York.

- Taylor, J. R. (1997). *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, University Science Books, Sausalito, California.

- Thomopoulos, S. C. A. (1989). "Sensor Integration and Data Fusion." *SPIE Vol. 1198: Sensor Fusion II: Human and Machine Strategies*, 178-191.

- Uzan, J., Scullion, T., Michalek, C. H., and Parades, M., Lytton, R. L. (1988). "A Microcomputer Based Procedure for Backcalculating Layer Moduli from FWD Data." *Texas Transportation Institute: Research Report No. 1123-1*, Texas A&M University, College Station, Texas.

- Yuan, D. and Nazarian, S. 1993. Automated Surface Wave Testing: Inversion Technique. Journal of Geotechnical Engineering, ASCE, 119(7), 1112-1126.

- Zadeh, L. A. (1965). "Fuzzy Sets." *Journal of Information and Control 8*, 338-353.

APPENDIX A

SUPPLEMENTAL TABLES FOR COMPARISON OF FUSION METHODS

In Chapter 8 graphs were shown comparing the different fusion methods, however, a point by point tabular comparison was not included. The tables in this appendix include combined result information from the applied fusion methods.

LIST OF TABLES FOR APPENDIX A

Table A.1 - TxDOT AC Modulus Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM
1	659	657	659	647	647	658	654	606	500	647	672
2	634	634	634	636	632	634	634	589	500	632	636
3	644	644	644	636	636	644	642	596	500	636	652
4	525	534	528	556	556	544	544	519	500	556	501
5	520	527	522	547	547	536	536	514	500	496	547
6	629	629	627	605	605	630	618	588	500	659	605
7	704	681	702	681	681	697	687	638	500	732	681
8	527	532	529	549	549	538	538	519	500	508	549
9	711	689	709	689	689	702	693	641	500	689	734
10	604	604	604	619	590	606	601	570	500	590	619
11	408	408	408	412	412	409	409	439	500	404	412
12	601	602	600	621	584	604	598	568	500	621	584
13	547	551	549	571	571	557	557	533	500	571	527
14	472	475	473	490	490	483	483	482	500	455	490
15	605	605	605	601	601	605	605	570	500	609	601
16	502	506	503	521	521	512	512	502	500	521	486
17	543	547	545	563	563	551	551	530	500	563	527
18	538	540	539	552	552	542	542	526	500	552	526
19	570	573	571	603	603	584	578	549	500	543	603
20	661	661	661	655	655	661	661	608	500	668	655
Avg	580	580	581	588	584	585	582	554	500	583	580
StDev	79	73	78	67	66	74	71	52	0	82	80
CoV	14%	13%	13%	11%	11%	13%	12%	9%	0%	14%	14%
Min	408	408	408	412	412	409	409	439	500	404	412
Max	711	689	709	689	689	702	693	641	500	732	734

Table A.2 - TxDOT Base Modulus Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM
1	63	82	63	68	68	83	83	77	68	102	60
2	61	61	62	68	68	62	62	58	43	62	68
3	84	106	85	92	92	99	99	95	54	139	92
4	63	60	64	73	73	68	68	58	55	45	73
5	74	89	75	72	72	82	82	84	73	107	72
6	110	130	109	105	119	125	126	135	119	182	105
7	100	131	98	93	93	114	105	138	152	169	93
8	97	147	95	91	91	103	92	163	156	242	91
9	99	137	97	92	92	109	99	154	189	180	92
10	84	104	84	80	80	94	91	100	89	132	80
11	61	68	61	81	81	66	66	63	43	81	64
12	98	113	98	99	78	106	102	108	78	146	99
13	87	103	87	82	82	96	93	104	112	117	82
14	93	115	92	92	92	103	94	112	77	167	92
15	107	114	107	104	127	111	118	114	110	127	104
16	88	102	88	92	92	95	97	93	62	126	92
17	64	64	65	72	72	74	74	65	78	44	72
18	66	63	66	69	69	65	65	62	59	57	69
19	63	81	63	103	103	83	83	71	47	103	64
20	104	107	104	102	113	105	109	108	108	113	102
Avg	83	99	83	87	88	92	90	98	89	122	83
StDev	17	27	17	13	17	18	18	32	41	51	15
CoV	21%	27%	20%	15%	19%	20%	20%	33%	46%	42%	18%
Min	61	60	61	68	68	62	62	58	43	44	60
Max	110	147	109	105	127	125	126	163	189	242	105

Table A.3 - TxDOT Subgrade Modulus Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM
1	15	16	14	20	20	18	20	18	21	20	12
2	15	17	14	20	20	17	20	18	22	20	12
3	13	15	13	11	19	14	16	16	19	17	11
4	12	15	12	10	10	14	13	16	17	21	10
5	13	15	12	10	10	15	13	16	18	20	10
6	13	15	12	10	10	15	12	17	19	21	10
7	14	16	13	11	11	15	16	17	19	21	11
8	13	14	13	11	19	14	15	15	19	15	11
9	12	13	12	10	10	12	12	13	17	13	10
10	11	13	11	17	17	13	13	14	16	17	9
11	10	12	10	16	16	15	15	13	14	16	8
12	11	14	11	21	9	16	13	15	16	21	9
13	13	15	12	10	10	14	13	16	18	20	10
14	12	13	12	10	10	12	12	14	17	14	10
15	12	13	12	10	10	12	12	14	17	14	10
16	11	14	11	21	17	16	14	16	17	21	9
17	11	14	11	21	17	16	14	16	17	21	9
18	11	13	11	17	17	14	14	14	17	17	9
19	11	13	11	17	17	14	14	14	16	17	9
20	12	13	12	10	10	13	13	14	18	14	10
Avg	12	14	12	14	14	14	14	15	18	18	10
StDev	1	1	1	5	4	2	2	1	2	3	1
CoV	11%	9%	9%	33%	30%	11%	16%	10%	10%	16%	11%
Min	10	12	10	10	9	12	12	13	14	13	8
Max	15	17	14	21	20	18	20	18	22	21	12

Table A.4 - TxDOT AC Uncertainty Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM
1	66	66	73	91	91	65	70	93	N/A	91	94
2	63	62	68	89	88	62	63	77	N/A	88	89
3	64	63	70	89	89	63	65	84	N/A	89	91
4	53	54	57	78	78	63	63	32	N/A	78	70
5	52	53	57	77	77	61	61	28	N/A	69	77
6	63	62	67	85	85	62	67	81	N/A	92	85
7	70	94	83	95	95	72	84	122	N/A	102	95
8	53	53	57	77	77	59	59	26	N/A	71	77
9	71	97	85	96	96	75	88	124	N/A	96	103
10	60	59	62	87	83	60	61	62	N/A	83	87
11	41	40	54	58	58	41	41	53	N/A	57	58
12	60	59	61	87	82	60	60	62	N/A	87	82
13	55	55	57	80	80	59	59	36	N/A	80	74
14	47	47	57	69	69	55	55	24	N/A	64	69
15	61	60	62	84	84	60	59	61	N/A	85	84
16	50	51	57	73	73	57	57	18	N/A	73	68
17	54	54	57	79	79	58	58	32	N/A	79	74
18	54	54	57	77	77	55	55	26	N/A	77	74
19	57	56	59	84	84	62	58	52	N/A	76	84
20	66	65	73	92	92	65	65	93	N/A	94	92
Avg	58	60	64	82	82	61	62	59	N/A	82	81
StDev	8	14	9	9	9	7	10	33	N/A	11	11
CoV	14%	23%	14%	11%	11%	11%	16%	55%	N/A	14%	14%
Min	41	40	54	58	58	41	41	18	N/A	57	58
Max	71	97	85	96	96	75	88	124	N/A	102	103

Table A.5 - TxDOT Base Uncertainty Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM
1	9	22	11	31	31	24	24	22	31	43	11
2	9	12	11	12	12	10	10	13	20	26	12
3	13	25	14	17	17	17	18	43	25	58	17
4	10	10	11	13	13	11	11	14	25	19	13
5	11	22	12	13	13	17	17	20	34	45	13
6	17	27	17	19	55	25	36	41	55	76	19
7	16	29	16	17	17	19	17	40	70	71	17
8	16	33	16	16	16	16	16	76	72	102	16
9	16	30	16	17	17	18	16	54	87	76	17
10	13	25	14	14	14	17	16	28	41	55	14
11	9	16	10	34	34	15	15	19	20	34	12
12	15	25	16	18	36	18	24	35	36	61	18
13	13	24	14	15	15	18	15	19	52	49	15
14	14	27	15	17	17	17	16	48	35	70	17
15	16	24	17	19	53	21	33	12	51	53	19
16	14	24	14	17	17	15	19	32	29	53	17
17	10	12	11	13	13	17	17	18	36	18	13
18	10	12	11	12	12	12	12	6	27	24	12
19	9	22	11	43	43	25	25	29	22	43	12
20	16	22	17	18	47	24	29	6	50	47	18
Avg	13	22	14	19	25	18	19	29	41	51	15
StDev	3	6	2	8	15	4	7	18	19	21	3
CoV	21%	29%	18%	43%	60%	24%	37%	62%	46%	42%	18%
Min	9	10	10	12	12	10	10	6	20	18	11
Max	17	33	17	43	55	25	36	76	87	102	19

Table A.6 - TxDOT Subgrade Uncertainty Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM
1	1.1	1.0	1.1	3.2	3.2	2.0	3.0	4.9	2.1	3.2	1.3
2	1.1	1.0	1.2	3.2	3.2	2.0	3.0	5.3	2.2	3.2	1.3
3	0.9	1.0	1.0	1.2	1.9	1.0	1.0	4.2	1.9	2.7	1.2
4	0.8	1.0	0.9	1.1	1.1	1.0	1.0	5.6	1.7	3.4	1.1
5	0.9	1.0	1.0	1.1	1.1	1.0	1.0	5.3	1.8	3.2	1.1
6	0.9	1.0	1.0	1.1	1.1	1.0	1.0	5.9	1.9	3.4	1.1
7	1.0	1.0	1.1	1.2	1.2	1.0	1.0	5.3	1.9	3.4	1.2
8	0.9	1.0	1.0	1.2	1.9	1.0	1.0	4.0	1.9	2.4	1.2
9	0.8	1.0	0.9	1.1	1.1	1.0	1.0	3.5	1.7	2.1	1.1
10	0.8	1.0	0.8	2.7	2.7	1.0	1.0	4.4	1.6	2.7	1.0
11	0.7	1.0	0.8	2.6	2.6	2.0	2.0	4.2	1.4	2.6	0.9
12	0.8	1.0	0.9	3.4	1.0	2.0	1.0	6.0	1.6	3.4	1.0
13	0.9	1.0	1.0	1.1	1.1	1.0	1.0	5.3	1.8	3.2	1.1
14	0.8	1.0	0.9	1.1	1.1	1.0	1.0	3.5	1.7	2.2	1.1
15	0.8	1.0	0.9	1.1	1.1	1.0	1.0	3.5	1.7	2.2	1.1
16	0.8	1.0	0.9	3.4	1.7	1.0	1.0	6.1	1.7	3.4	1.0
17	0.8	1.0	0.9	3.4	1.7	2.0	1.0	6.1	1.7	3.4	1.0
18	0.8	1.0	0.9	1.7	1.7	1.0	1.0	4.6	1.7	2.7	1.0
19	0.8	1.0	0.8	2.7	2.7	1.0	1.0	4.4	1.6	2.7	1.0
20	0.8	1.0	0.9	1.1	1.1	1.0	1.0	4.0	1.8	2.2	1.1
Avg	0.9	1.0	0.9	1.9	1.7	1.3	1.3	4.8	1.8	2.9	1.1
StDev	0.1	0.0	0.1	1.0	0.8	0.4	0.6	0.9	0.2	0.5	0.1
CoV	11%	0%	11%	51%	44%	36%	51%	19%	10%	16%	11%
Min	0.7	1.0	0.8	1.1	1.0	1.0	1.0	3.5	1.4	2.1	0.9
Max	1.1	1.0	1.2	3.4	3.2	2.0	3.0	6.1	2.2	3.4	1.3

Table A.7 - TTI AC Modulus Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM	PSPA
1	498	633	490	642	471	563	483	609	500	642	821	471
2	486	513	485	564	502	524	510	510	500	564	502	473
3	494	546	491	480	480	530	502	540	500	521	659	480
4	510	543	507	495	495	534	516	533	500	583	555	495
5	529	545	526	518	518	537	526	535	500	604	519	518
6	464	484	464	525	525	501	496	487	500	464	525	459
7	510	517	510	521	521	530	522	512	500	452	575	521
8	481	469	482	489	489	486	486	472	500	497	401	489
9	513	512	515	528	528	544	535	498	500	593	371	528
10	546	533	545	550	486	543	526	525	500	564	486	550
11	489	519	488	481	481	518	501	515	500	501	577	481
12	484	499	484	478	478	497	496	499	500	501	516	478
13	454	475	454	515	515	494	494	480	500	515	460	444
14	464	485	464	529	529	504	498	488	500	464	529	459
15	525	508	525	535	503	521	505	506	500	503	487	535
16	492	463	493	523	523	513	498	468	500	427	421	523
17	490	471	490	503	503	493	492	477	500	456	449	503
18	478	456	483	506	506	500	498	457	500	468	352	506
19	546	567	541	542	434	532	481	554	500	740	434	542
20	536	504	534	559	497	517	482	502	500	497	452	559
Avg	499	512	498	524	499	519	502	508	500	528	505	501
StDev	27	42	26	38	24	20	16	35	0	76	105	33
CoV	5%	8%	5%	7%	5%	4%	3%	7%	0%	14%	21%	7%
Min	454	456	454	478	434	486	481	457	500	427	352	444
Max	546	633	545	642	529	563	535	609	500	740	821	559

Table A.8 - TTI Base Modulus Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM
1	38	70	39	44	44	67	58	74	147	44	31
2	69	87	67	64	69	80	91	99	164	64	69
3	55	117	54	58	48	55	52	142	320	58	48
4	49	96	50	42	42	76	48	110	211	42	76
5	55	80	55	47	47	65	62	81	125	47	72
6	61	84	62	46	46	75	70	87	135	81	46
7	43	72	43	52	52	63	61	74	132	38	52
8	56	81	57	47	47	56	53	79	96	47	95
9	49	79	49	46	46	65	54	89	170	46	50
10	60	71	60	60	53	60	60	70	97	60	53
11	65	78	64	56	56	67	70	80	116	67	56
12	50	75	51	62	44	61	60	74	117	44	62
13	56	72	57	49	49	59	57	71	98	49	67
14	45	59	45	43	43	46	46	56	83	43	42
15	47	74	48	56	42	63	61	78	136	42	56
16	43	54	44	40	40	44	44	51	70	40	44
17	28	55	29	35	35	48	48	46	80	24	35
18	30	65	32	41	41	60	60	58	106	26	41
19	29	48	33	46	46	46	46	43	61	46	21
20	27	39	31	52	52	51	51	35	33	52	20
Avg	48	73	49	49	47	60	58	75	125	48	52
StDev	12	17	11	8	7	10	11	25	62	13	19
CoV	26%	24%	24%	16%	15%	17%	19%	33%	49%	28%	36%
Min	27	39	29	35	35	44	44	35	33	24	20
Max	69	117	67	64	69	80	91	142	320	81	95

Table A.9 - TTI Subgrade Modulus Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM
1	10	12	11	14	14	12	12	12	14	12	9
2	11	20	12	28	28	22	22	17	14	28	10
3	13	18	14	16	16	15	15	16	16	22	11
4	13	29	14	16	16	24	15	26	16	50	11
5	14	26	15	12	12	20	16	24	17	44	12
6	15	31	16	13	13	17	16	30	19	59	13
7	16	26	17	14	14	19	17	26	20	43	14
8	18	21	18	16	16	19	20	21	23	24	16
9	17	20	17	15	15	18	18	20	22	24	15
10	17	28	17	15	15	20	18	29	23	49	15
11	16	28	17	14	14	19	17	28	21	48	14
12	19	19	19	17	24	20	20	19	24	16	17
13	18	24	18	16	16	20	24	25	23	35	16
14	14	25	15	12	12	19	17	23	17	40	12
15	16	25	17	14	14	19	18	24	20	39	14
16	15	26	16	13	13	19	17	25	19	42	13
17	18	20	18	16	16	19	19	20	23	21	16
18	19	20	19	17	17	20	20	20	25	19	17
19	9	13	11	13	13	13	13	11	16	4	13
20	12	19	13	11	11	16	16	17	14	26	11
Avg	15	23	16	15	15	19	18	22	19	32	13
StDev	3	5	2	4	4	3	3	5	4	15	2
CoV	20%	23%	16%	23%	26%	15%	16%	24%	19%	46%	17%
Min	9	12	11	11	11	12	12	11	14	4	9
Max	19	31	19	28	28	24	24	30	25	59	17

Table A.10 - TTI AC Uncertainty Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM	PSPA
1	29	61	30	93	31	50	29	160	N/A	90	172	31
2	28	46	29	82	105	49	47	39	N/A	79	105	31
3	28	49	29	32	32	41	39	81	N/A	73	138	32
4	29	48	30	33	33	39	31	43	N/A	82	117	33
5	30	48	32	34	34	34	40	47	N/A	85	109	34
6	27	45	27	109	109	70	63	31	N/A	65	110	30
7	29	46	30	34	34	37	34	51	N/A	63	121	34
8	28	37	28	32	32	29	29	47	N/A	70	84	32
9	30	39	32	35	35	35	33	93	N/A	83	78	35
10	31	45	34	36	101	39	61	38	N/A	79	102	36
11	28	47	29	32	32	46	34	43	N/A	70	121	32
12	28	45	28	31	31	43	41	16	N/A	70	108	31
13	26	42	27	74	74	57	57	33	N/A	72	97	29
14	27	45	27	110	110	71	62	33	N/A	65	111	30
15	30	43	31	35	73	32	48	20	N/A	70	102	35
16	28	35	29	34	34	32	31	51	N/A	60	88	34
17	28	38	28	33	33	29	29	28	N/A	64	94	33
18	28	33	28	33	33	30	30	72	N/A	66	74	33
19	32	45	34	36	91	33	54	132	N/A	104	91	36
20	31	41	33	37	72	35	55	44	N/A	70	95	37
Avg	29	44	30	49	56	41	42	55	N/A	74	106	33
StDev	2	6	2	28	31	12	12	37	N/A	11	22	2
CoV	5%	14%	7%	57%	56%	30%	29%	66%	N/A	14%	21%	7%
Min	26	33	27	31	31	29	29	16	N/A	60	74	29
Max	32	61	34	110	110	71	63	160	N/A	104	172	37

Table A.11 - TTI Base Uncertainty Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	MODULUS	SMART	JIM
1	8	22	10	12	12	19	14	64	72	12	11
2	14	21	15	18	25	17	24	56	80	18	25
3	12	39	13	16	17	12	12	154	157	16	17
4	11	27	12	12	12	18	11	89	103	12	27
5	12	21	13	13	13	14	14	40	61	13	26
6	13	20	14	17	17	22	21	45	66	23	17
7	9	21	11	19	19	16	15	51	65	11	19
8	12	21	13	13	13	12	13	28	47	13	34
9	10	23	12	13	13	15	11	70	83	13	18
10	12	19	13	17	19	12	14	24	48	17	19
11	13	20	14	20	20	14	19	32	57	19	20
12	11	21	12	22	12	13	13	38	57	12	22
13	11	19	13	14	14	12	12	25	48	14	24
14	9	18	10	12	12	9	9	23	41	12	15
15	10	21	11	20	12	15	13	51	67	12	20
16	9	15	10	11	11	9	9	16	34	11	16
17	6	20	8	13	13	14	14	30	39	7	13
18	6	23	9	15	15	19	19	43	52	7	15
19	6	14	9	13	13	12	12	20	30	13	8
20	6	9	8	15	15	14	14	16	16	15	7
Avg	10	21	12	15	15	14	14	46	61	13	19
StDev	3	6	2	3	4	3	4	32	30	4	7
CoV	25%	28%	18%	21%	24%	23%	27%	70%	49%	28%	36%
Min	6	9	8	11	11	9	9	16	16	7	7
Max	14	39	15	22	25	22	24	154	157	23	34

Table A.12 - TTI Subgrade Uncertainty Comparison

Test Pt.	Stat. Wt. Avg.	Fuzzy	Hybrid	Cum. Decision	D-T Decision	Cum. Fusion	D-T Fusion	Average	Mod 5.1	SMART	JIM
1	1.3	2.0	2.0	2.7	2.7	2.0	2.0	2.5	2.7	5.5	1.5
2	1.4	6.0	2.0	12.9	12.9	8.0	8.0	9.5	2.7	12.9	1.7
3	1.7	5.0	2.0	3.0	3.0	2.0	2.0	5.5	3.0	10.1	1.9
4	1.7	9.0	2.0	3.0	3.0	6.0	2.0	21.2	3.0	23.0	1.9
5	1.8	8.0	2.0	2.0	2.0	4.0	2.0	17.2	3.2	20.2	2.0
6	2.0	9.0	2.0	2.2	2.2	2.0	2.0	25.0	3.6	27.1	2.2
7	2.1	7.0	2.0	2.4	2.4	3.0	2.0	15.3	3.8	19.8	2.4
8	2.3	4.0	3.0	2.7	2.7	2.0	3.0	4.4	4.4	11.0	2.7
9	2.2	4.0	2.0	2.6	2.6	2.0	2.0	4.7	4.2	11.0	2.6
10	2.2	7.0	3.0	2.6	2.6	3.0	2.0	17.8	4.4	22.5	2.6
11	2.1	7.0	2.0	2.4	2.4	3.0	2.0	18.0	4.0	22.1	2.4
12	2.3	3.0	3.0	2.9	4.6	3.0	3.0	4.4	4.6	7.4	2.9
13	2.3	5.0	3.0	2.7	2.7	3.0	6.0	9.6	4.4	16.1	2.7
14	1.8	7.0	2.0	2.0	2.0	4.0	2.0	14.9	3.2	18.4	2.0
15	2.1	6.0	2.0	2.4	2.4	3.0	2.0	13.1	3.8	17.9	2.4
16	2.0	7.0	2.0	2.2	2.2	3.0	2.0	15.3	3.6	19.3	2.2
17	2.3	4.0	3.0	2.7	2.7	2.0	3.0	3.6	4.4	9.7	2.7
18	2.3	4.0	3.0	2.9	2.9	2.0	4.0	4.2	4.8	8.7	2.9
19	1.3	2.0	2.0	2.2	2.2	2.0	2.0	6.2	3.0	1.8	2.2
20	1.6	6.0	2.0	1.9	1.9	4.0	4.0	7.9	2.7	12.0	1.9
Avg	1.9	5.6	2.3	3.0	3.1	3.2	2.9	11.0	3.7	14.8	2.3
StDev	0.4	2.1	0.5	2.3	2.4	1.5	1.6	6.7	0.7	6.8	0.4
CoV	18%	37%	20%	78%	76%	49%	56%	61%	19%	46%	17%
Min	1.3	2.0	2.0	1.9	1.9	2.0	2.0	2.5	2.7	1.8	1.5
Max	2.3	9.0	3.0	12.9	12.9	8.0	8.0	25.0	4.8	27.1	2.9