2002-25

Final Report

Delineation of the Stiff Layer From FWD Measurements

Research

Technical Report Documentation Page

1. Report No.	2.	3. Recipients Accession No.
MN/RC - 2002-25		

4. Title and Subtitle	5. Report Date
Delineation of the Stiff Layer from FWD Measurements	October 2001
	6.

7. Author(s)	8. Performing Organization Report No.
Bojan Guzina, Dongwei Cao	

9. Performing Organization Name and Address	10. Project/Task/Work Unit No.
University of Minnesota Department of Civil Engineering 122 Civil Engineering 500 Pillsbury Drive SE Minneapolis, MN 55455	11. Contract (C) or Grant (G) No. (c) 74708 (wo) 137

12. Sponsoring Organization Name and Address	13. Type of Report and Period Covered
Minnesota Department of Transportation 395 John Ireland Boulevard Mail Stop 330 St. Paul, Minnesota 55155	Final 2001 14. Sponsoring Agency Code

15. Supplementary Notes

http://www.lrrb.org/PDF/200225.pdf

Please note: The blurriness of the screen version of this report is not present when the report is printed to hard copy. Sorry for any inconvenience.

16. Abstract (Limit: 200 words)

The Falling Weight Deflectometer (FWD) is a widely used non-destructive test device for estimating the pavement stiffness properties. However, the conventional elastostatic interpretation of FWD measurements is generally associated with a number of inconsistencies. The purpose of this project is to develop a reliable and effective dynamic backcalculation method capable of estimating the location and properties of the permanent or seasonal stiff layer (as well as other pavement stiffness properties) from FWD measurements. The backcalculation method is implemented in the form of a user-friendly software that allows unedited deflection time histories from the FWD test to be used as an input to the back-analysis. The backcalculation scheme developed in this study is based on the Artificial Neural Network (ANN) approach and employs a three-dimensional multilayer viscoelastic dynamic model as a predictive tool.

17. Document Analysis/Descriptors		18. Availability Statement
Falling Weight Deflectometer	FWD Testing	No restrictions. Document available from:
Backcalculation	Artificial Neural Network	National Technical Information Services,
Hot mix asphalt	pavement	Springfield, Virginia 22161

19. Security Class (this report)	20. Security Class (this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	124	

Delineation of the Stiff Layer from FWD Measurements

Final Report

Prepared by

Bojan B. Guzina
Assistant Professor
Department of Civil Engineering
University of Minnesota

and

Dongwei Cao
Graduate Student
Department of Civil Engineering
University of Minnesota

October 2001

Prepared for the

Minnesota Department of Transportation
Office of Materials and Road Research
1400 Gervais Avenue, Maplewood, MN 55109-2044

Acknowledgement

The support of the Mn/DOT Office of Materials and Road Research during the course of this investigation is gratefully acknowledged. The authors would especially like to express their gratitude to the following individuals:

- John Siekmeier, P.E., Office of Materials and Road Research, Mn/DOT.
- Bruce A. Chadbourn, Office of Materials and Road Research, Mn/DOT.
- Peter Davich, Office of Materials and Road Research, Mn/DOT.
- Professor Andrew Drescher, Department of Civil Engineering, University of Minnesota.
- Professor Mihai Marasteanu, Department of Civil Engineering, University of Minnesota.
- Sylvain Nintcheu, Department of Civil Engineering, University of Minnesota.
- Olivier Hoffmann, Department of Civil Engineering, University of Minnesota.

Contents

List of Figures

List of Tables

Executive Summary

The Falling Weight Deflectometer (FWD) is a widely used non-destructive test device for estimating the pavement stiffness properties. However, the conventional elastostatic interpretation of FWD measurements is generally associated with a number of inconsistencies (e.g. Chang et.al. [34]). The purpose of this project is to develop a reliable and effective dynamic backcalculation method capable of estimating the location and properties of the permanent or seasonal stiff layer (as well as other pavement stiffness properties) from FWD measurements. The backcalculation method is implemented in the form of a user-friendly software that allows unedited deflection time histories from the FWD test to be used as an input to the back-analysis.

The backcalculation scheme developed in this study is based on the Artificial Neural Network (ANN) approach and employs a three-dimensional multilayer viscoelastic dynamic model (Guzina and Pak, [4]) as a predictive tool. The key features of the proposed backcalculation scheme are:

- The hot mix asphalt (HMA) is characterized as a viscoelastic material using the power law model (Findley, [12]) whose coefficients are assumed to be known and are used as an *input* to the ANN-based backcalculation. For a specific type of the hot mix asphalt, these parameters are calibrated via an empirical formula (Witczak, [17]) describing the HMA's stiffness variation with frequency as a function of the mix characteristics and HMA temperature.

- The FWD measurements are interpreted in terms of the pavement's frequency response functions (FRF's), which are estimated from the measured FWD deflection time histories via the discrete Fourier transform.

- Due to the finite width of the pavement profile, wave reflections from lateral pavement edges are found to influence the FWD measurements significantly, which is a phenomenon that is not accounted for by the elastodynamic model used. To minimize the effect of side reflections, the useful frequency range for the frequency

response functions that are used as an input to the back-analysis is taken to be 1 Hz $\leq f \leq$ 51 Hz, thus avoiding the pronounced resonance peak due to the lateral wave reflections typically located between 50 Hz and 70 Hz.

- The inputs to the featured ANN backcalculation scheme are (i) the frequency response functions synthesizing the FWD measurements, (ii) the coefficients of the power law model describing the HMA properties, and (iii) the thicknesses of the HMA and aggregate base layers. The outputs of the backcalculation scheme are the shear moduli of aggregate base, subgrade, and stiff layer (i.e. bottom half-space), as well as the thickness of the subgrade which represents the depth to the stiff layer.

The proposed ANN backcalculation scheme is applied to several field FWD measurements. The predicted pavement properties are compared to those given by the conventional static backcalculation software Evercalc ([33]) for a representative field test. Table 1 shows the comparison where "E" stands for the Young's modulus, "h" denotes the layer thickness, while subscripts "B", "S" and "H" stand for the aggregate base, subgrade, and half-space (i.e. stiff layer), respectively. From the Table, it can be seen that Evercalc yields a smaller value for the stiffness of the aggregate base (E_B) than that of the subgrade (E_S), which is contrary to the physical situation. In contrast, the ANN backcalculation scheme (row labeled "ANN" in the Table) gives a reasonable estimate, i.e. $E_B > E_S$. Another advantage of the ANN backcalculation scheme over Evercalc is that *both* depth to the stiff layer and its stiffness are given as outputs, while either of them must be specified in advance in Evercalc. As for the thickness of the subgrade (i.e. depth to the stiff layer), Evercalc apparently overestimates its true value because the prediction method used by Evercalc is based on the assumption that the stiff layer is approximately 100 times stiffer than the subgrade ([33]).

To improve the performance of the ANN backcalculation model developed, the following recommendations are made:

- The mechanical properties of the HMA layer are characterized by the power law

Table 1: ANN vs. Evercalc (Test Section 33, May 22, 2001)

Parameter	E_B [MPa]	E_S [MPa]	h_S [m]	E_H [MPa]
ANN	136	87	0.9	265
Evercalc	44	151	2.8	176

model. Currently, the coefficients of the power law model are obtained based on the empirical equation given by Witczak ([17]) describing HMA's Young's modulus as a function of the HMA aggregate coefficients, temperature and the viscosity of the asphalt binder. A method capable of producing more accurate estimates of the power law coefficients would certainly be beneficial.

- The FWD deflection records, which are inherently affected by the noise accumulated through integrating the geophone (i.e. velocity) measurements, are represented by the suitable frequency response functions in the backcalculation scheme. A significant difference between the FRF estimates based on various baseline correction methods (i.e. methods designed to delete the noise accumulated in the FWD deflection records) is observed, which has a substantial effect on the estimated pavement properties. Therefore, a more accurate baseline correction method is expected to further improve the backcalculation performance. On the other hand, according to the definition of the frequency response function and Fourier transform, the FRF's of interest in this study can be obtained through either deflection, velocity, or acceleration time histories. Therefore, a *direct* use of the *velocity records* which do not require the baseline correction (but are currently unavailable through the FWD setup) would substantially improve the prediction quality of the backcalculation scheme proposed.

- To further alleviate the problem associated with lateral reflections in FWD testing, it is recommended that the FWD test be performed near the center line of the pavement (i.e. away from pavement edges).

- The ANN backcalculation scheme is designed to interpret the deflection measurements at distances 0.305, 0.914 and 1.829 meters from the center of the loading plate. To allow for the continuing use of the backcalculation method developed, it is recommended that the geophones at these distances be retained in future alterations of the FWD setup.

1 Introduction

The Falling Weight Deflectometer (FWD) is a non-destructive testing device widely used in pavement assessment. As shown in Figs. 1 and 2, the FWD test is performed by applying an impulse load on the pavement surface through a circular loading plate and monitoring the induced pavement deflection at several distances from the source. In the test, the recorded impact force and displacement signals are used as a basis for estimating the pavement stiffness properties, which is an inverse problem of interest in this study. There are two major parts in solving the inverse problem: one is the choice of an appropriate forward model that can be used to simulate the FWD test, and the other is a suitable inversion algorithm.

With regard to the first problem component, the forward model that simulates the FWD test is typically given by an elastostatic multi-layered model (e.g. Khazanovich and Roesler, [1]), in which the equivalent static load on the pavement is represented by the peak value of the FWD load time history, while the equivalent static displacement is represented by the peak value of the FWD deflection record. More realistically, the FWD test can be characterized via an elastodynamic model (e.g. Meier and Rix, [2]) in which the entire load time history is used to calculate the transient pavement deflection at control points. Because the FWD test is essentially a dynamic test and the dynamic backcalculation methods are typically able to extract more information from FWD measurements (Magnuson et.al [3]), an interpretation method based on elastodynamic theory is preferred over that based on static models. Consequently, in the current project, the multilayer viscoelastic dynamic model developed by Guzina & Pak ([4]) is used as a basis for the forward analysis.

Given the forward model (e.g. an elastodynamic multilayered model), various techniques are available to solve the inverse problem. Most available methods are based on the iterative gradient-based minimization of an objective function (Harichandran, et.al. [5], Hossain and Zaniewski [6]). In this class of methods, an assumed objective (i.e. cost) function is usually defined as a measure of the misfit between the theoretical and

measured deflections at control locations. When the objective function is minimized, the pavement properties used in calculating the theoretical deflections at the minimum point are taken as the backcalculation result. For example, Harichandran et. al [5] employed the modified Newton algorithm to backcalculate the pavement properties. With the development of artificial intelligence, an alternative minimization method known as Genetic Algorithm (Fwa et. al. [7]) has also been employed to solve the foregoing class of inverse problems. However, the main drawback of the gradient-based minimization methods and genetic algorithms is that they require an extensive computation time because the forward (i.e. predictive) model needs to be run multiple times. The artificial neural network (ANN) approach overcomes this problem because it eliminates the need to evaluate the forward model during backcalculation of the pavement properties from FWD measurements (Kim and Kim [8]; Meier and Rix [2]; Khazanovich and Roesler [1]; Williams and Gucunski [9]).

In the current project, a novel backcalculation technique is developed to estimate the pavement stiffness properties from the FWD measurements based on a dynamic pavement model and an artificial neural network approach. The advantages of the proposed backcalculation model involve: (i) the use of a three-dimensional multilayer visco-elastodynamic model as a predictive tool that accounts for the dynamic nature of the FWD test and the creep behavior of the Hot Mix Asphalt (HMA) layer, (ii) the computation efficiency of the ANN, and (iii) the fact that the ANN predictions do not rely on initial seed values. In the report, the term "HMA" and "asphalt concrete" will be used interchangeably.

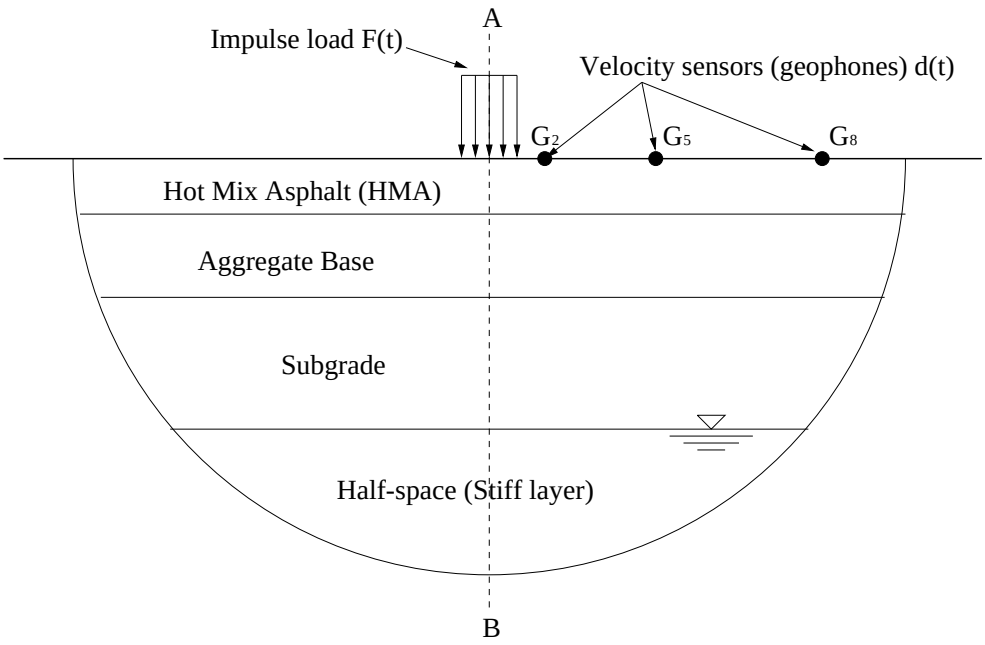

Figure 1: Schematics of the FWD (Falling Weight Deflectometer) Test

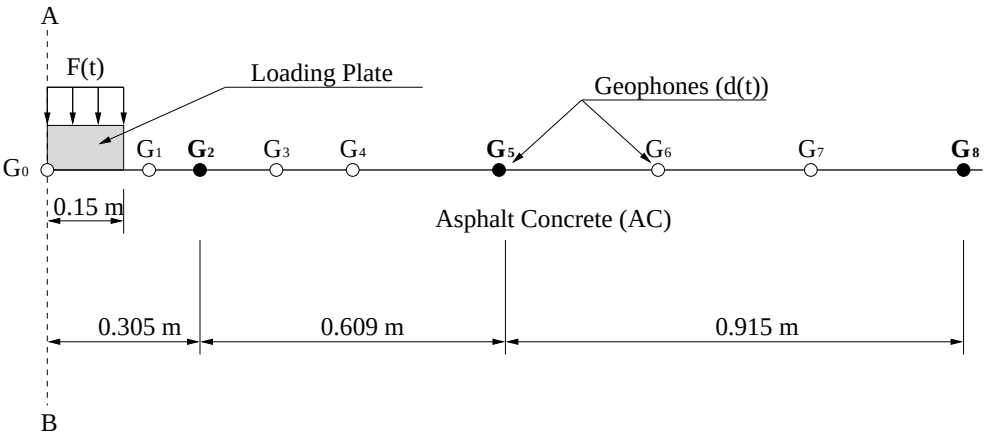

Figure 2: Layout of the FWD apparatus

2 Physical Problem

2.1 Experimental Setup

Fig. 2 shows the setup of the Falling Weight Deflectometer used by the Minnesota Department of Transportation (Mn/DOT). In the setup, a buffered annular loading plate with a radius of 15 cm resting on the pavement surface is used to transfer the impact load from the falling weight to the pavement. The induced vertical surface displacement is monitored at nine control points using the velocity sensors (G_i, $i = 0$, 1, ..., 8) known as geophones whose spacing is listed in Table 2. In the current project, the velocity measurements at geophones 2, 5 and 8 are used as a basis for the wave-based pavement characterization.

2.2 Back-analysis of Pavement Properties

As shown in Fig. 3, the pavement profile in this study is approximated as a multilayer visco-elastic system consisting of the hot mix asphalt (HMA) layer, aggregate base, subgrade, and semi-infinite half-space. The interface between the subgrade and half-space is taken to represent the location of a permanent or seasonal stiff layer (e.g. groundwater table or frozen ground), which is an important factor in determining the highway load restrictions. As such, the depth to the stiff layer will be one of the key issues addressed in the current project. Also of interest in this study are the stiffness (i.e. shear modulus) and thickness of each pavement layer. It should be noted that throughout this report, "E" stands for the Young's modulus, "G" indicates the shear modulus, "h" denotes the layer thickness, subscript "HMA" stands for the asphalt concrete, while subscripts "B", "S", and "H" refer to the aggregate base, subgrade, and half-space (i.e. stiff layer), respectively.

In the quest to resolve the elastic moduli of individual pavement layers from the FWD measurements, it should be kept in mind that the significant frequencies generated by the impulse load and recorded by the velocity sensors in the FWD test rarely exceed

80 to 100 Hz, so that the corresponding wave lengths (λ_{min}) in the asphalt layer are typically longer than 2 meters. On the basis of an empirical finding that the seismic waves are typically not capable of resolving layers thinner than $\lambda_{min}/12$ (Farr, [10]), it may be difficult to estimate properties of the HMA layer (whose thickness in the field is typically 0.1 m to 0.2 m) exclusively from FWD measurements. Moreover, the thickness and the basic material properties of the HMA layer are available from the construction record for many existing pavements. As a result, the thickness and the stiffness of the HMA layer are taken as *inputs* to the proposed backcalculation methodology. In this study, also assumed available (i.e. known) is the thickness of the aggregate base, which is taken as an *input* to the backcalculation model as well.

To summarize, the inputs for the backcalculation model developed in this study are (i) the thickness and material properties of HMA (i.e. asphalt concrete) layer, (ii) the thickness of the aggregate base, and (iii) the FWD force and deflection measurements; the outputs are (i) the stiffness of the aggregate base, subgrade and half-space, and (ii) the thickness of subgrade. The above categorization of the pavement properties and the basic structure of the proposed backcalculation approach are summarized in Table 3 and Fig. 4, respectively.

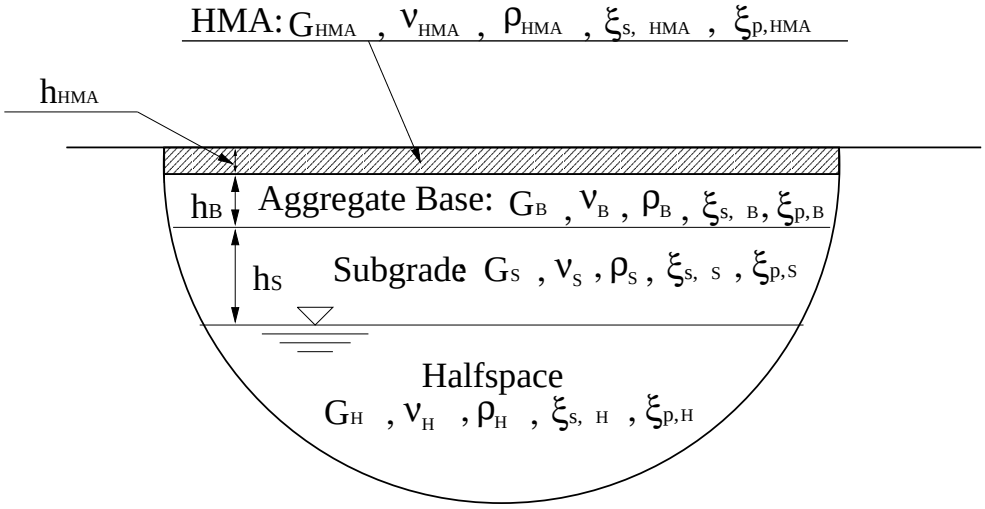

Figure 3: Pavement Profile

Table 2: Geophone Spacing in the FWD Setup

Geophone Number	Distance to center of loading plate [m]
0	0
1	0.203
2	**0.305**
3	0.457
4	0.610
5	**0.914**
6	1.219
7	1.524
8	**1.829**

Table 3: Classification of Pavement Properties

Layer	Known Properties	Unknown Properties
HMA	Stiffness, Thickness	(N/A)
Aggregate Base	Thickness	Stiffness
Subgrade	(N/A)	Thickness, Stiffness
Half-space	(N/A)	Stiffness

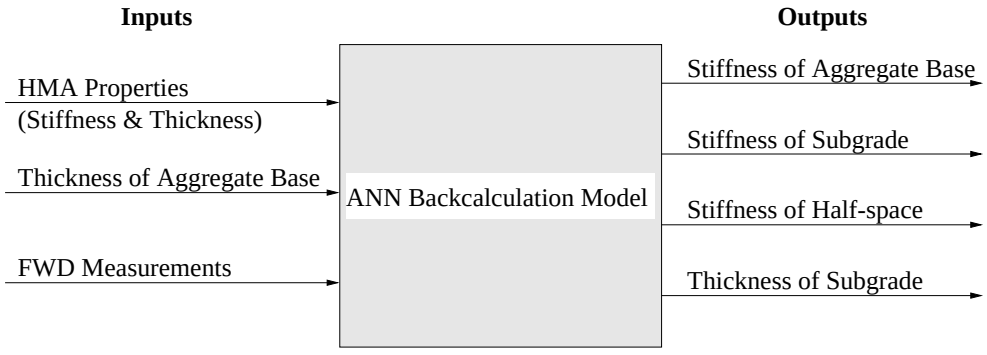

Figure 4: Overall Backcalculation Scheme

3 Forward Model

3.1 Introduction

In this investigation, the FWD test is simulated using the three-dimensional dynamic model developed in Guzina and Pak [4]. In this model, the pavement is represented by a multilayer viscoelastic system resting on a semi-infinite half-space (see Fig. 5). Each layer extends infinitely in the horizontal direction and is characterized by the shear modulus (G), Poisson's ratio (ν), mass density (ρ), and the damping ratios ξ_s and ξ_p for shear and compressional waves, respectively. The load acting on the pavement surface is assumed to be time harmonic and uniformly distributed over the circular contact area. This assumption is considered reasonable because, upon applying the Fourier transform, the impulse load in the FWD test can be decomposed into a series of time harmonic loads at different frequencies. Also, based on the study of Guzina and Nintcheu ([11]), the systematic errors induced by approximating the contact load distribution in the FWD test as uniform are insignificant as long as the deflection measurements are performed outside of the loaded annular area.

It should be noted that in the dynamic analysis of linear viscoelastic systems, material damping can be effectively taken into account by introducing the complex shear modulus (G^*) and the complex Poisson's ratio (ν^*) instead of the respective elastic constants G and ν into the linear elastic analysis (Findley et. al. [12]). In what follows, the choice of the complex shear modulus and Poisson's ratio for the asphalt concrete as well as the base, subgrade, and half-space soils will be discussed in detail with reference to Table 3 and the assumptions of above forward model.

3.2 Hot-Mix Asphalt (HMA) Layer

For the hot-mix asphalt (HMA) layer, the mass density (ρ_{HMA}) and the elastic Poisson's ratio (ν_{HMA}) do not vary significantly for different mixtures and are consequently assumed to be constant. On the other hand, the thickness of the HMA layer (h_{HMA})

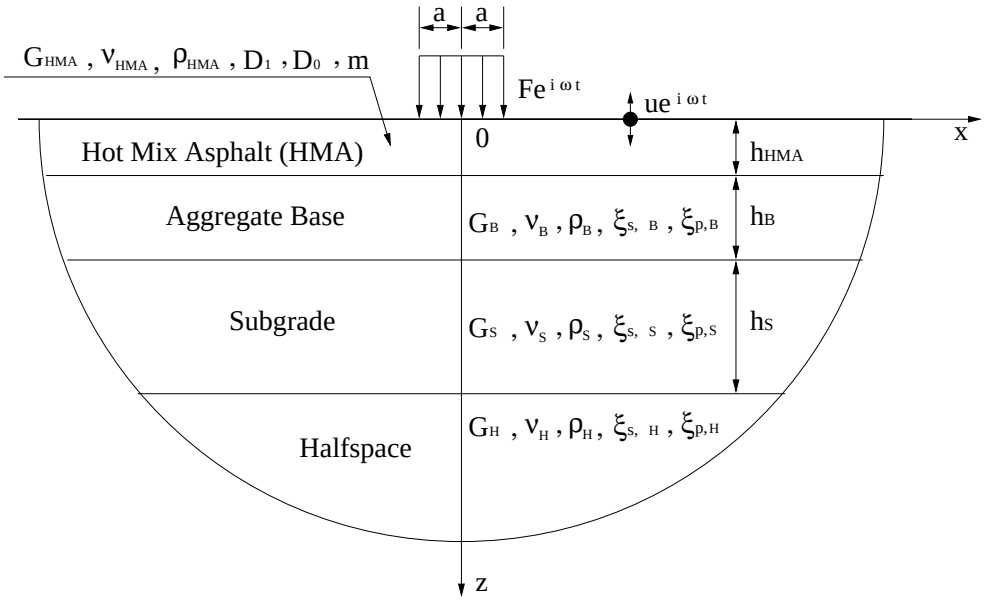

Figure 5: Forward Model

is assumed to be available from the construction record of the pavement tested. The particular values of the mass density, Poisson's ratio and the thickness of the asphalt concrete layer that are considered in this study are:

$$\rho_{HMA} = 2335 \text{ kg/m}^3, \qquad \nu_{HMA} = 0.4 \tag{1}$$

$$0.1 \text{ m} \leq h_{HMA} \leq 0.4 \text{ m} \qquad \text{(available from construction record)}$$

Owing to the asphalt's creep behavior, its complex shear (G^*_{HMA}) modulus and complex Poisson's ratio (ν^*_{HMA}) are in general frequency-dependent and a variety of experimental methods are available to measure their frequency variation. For generality of the neural network developed in this study, the complex shear modulus and complex Poisson's ratio of asphalt concrete that are used to simulate the FWD measurements are approximated using the power law model (Findley et. al. [12], Zhang and Drescher [13]).

According to the theory of viscoelasticity (Findley et. al. [12]), if the uniaxial creep compliance of a viscoelastic material in time domain ($J(t)$) is known, its strain-stress

relationship can be stated as:

$$\varepsilon(t) = \int_0^t J(t-\xi)\frac{\partial\sigma(\xi)}{\partial\xi} \tag{2}$$

By substituting $\xi' = t - \xi$ into eq. (2), one has:

$$\varepsilon(t) = -\int_0^\infty J(\xi')\frac{\partial(\sigma(t-\xi'))}{\partial\xi'} \tag{3}$$

Further, upon substituting $\varepsilon(t) = \varepsilon^* e^{i\omega t}$ and $\sigma(t) = \sigma_0 e^{i\omega t}$ (where ω is the circular frequency) into eq. (3), one obtains:

$$\varepsilon^* e^{i\omega t} = -\int_0^\infty J(\xi')\frac{\partial(\sigma_0 e^{i\omega(t-\xi')})}{\partial\xi'} \tag{4}$$

As a result, the complex uniaxial creep compliance $J^*(\omega)$ in the frequency domain can be expressed as:

$$J^* = \frac{\epsilon^*}{\sigma_0} = (i\omega)\int_0^\infty J(\xi')e^{-i\omega\xi'}d\xi'. \tag{5}$$

Owing to the fact that $J(\xi') = 0$ when $\xi' < 0$, eq. (5) can rewritten as:

$$J^* = (i\omega)\int_{-\infty}^\infty J(\xi')e^{-i\omega\xi'}d\xi' = (i\omega)\mathcal{F}[J(\xi')] \tag{6}$$

where $\mathcal{F}[J(\xi')]$ denotes the Fourier transform of $J(\xi')$ (see [15]).

In this investigation, the creep compliance behavior in time domain for the hot-mix asphalt is described by the power law model:

$$J(t) = D_0 + D_1 t^m \tag{7}$$

where D_0, D_1 and m are material parameters, with $1/D_0$ denoting the elastic Young's modulus. Substituting this relationship into (6), the complex compliance for the HMA can be written as

$$J^*(\omega) = D_0 + D_1\Gamma(1+m)(i\omega)^{-m} \tag{8}$$

By virtue of the fact that the uniaxial complex compliance and the complex Young's modulus (E^*) are reciprocal, i.e.

$$J^*(\omega)E^*_{HMA}(\omega) = 1, \tag{9}$$

The complex Young's modulus of the HMA can be obtained from (8) as

$$E^*_{HMA}(\omega) = \frac{1}{J^*(\omega)} = \frac{1}{D_0 + D_1\Gamma(1+m)(i\omega)^{-m}} \tag{10}$$

On the basis of (10) and an additional assumption that the complex bulk modulus of HMA (K^*_{HMA}) is equal to the elastic bulk modulus (K_{HMA}) (e.g. Hopman [16]), i.e.

$$K^*_{HMA} = K_{HMA} = \frac{E_{HMA}}{3(1-2\nu_{HMA})} \tag{11}$$

where $\nu_{HMA} = 0.4$ is the elastic Poisson's ratio and E_{HMA} is the elastic Young's modulus of asphalt concrete, the complex shear modulus (G^*_{HMA}) and complex Poisson's ratio (ν^*_{HMA}) of the asphalt concrete layer are finally obtained as:

$$G^*_{HMA} = \frac{3K_{HMA}E^*_{HMA}}{9K_{HMA} - E^*_{HMA}} \quad , \tag{12}$$

$$\nu^*_{HMA} = \frac{3K_{HMA} - E^*_{HMA}}{6K_{HMA}} \quad . \tag{13}$$

According to the physical meaning of power law model, the elastic Young's modulus (E_{HMA}) is the inverse of D_0, i.e.

$$E_{HMA} = \frac{1}{D_0} \tag{14}$$

As a result, the complex shear modulus (G^*_{HMA}) and complex Poisson's ratio (ν^*_{HMA}) of HMA can be expressed as a function of D_0, D_1 and m only. For the backcalculation model developed here, there are no restrictions on how to obtain these three parameters, as long as the values obtained characterize the HMA appropriately. In the current project, D_0, D_1 and m are obtained by minimizing the difference between the amplitude of the complex Young's modulus given by the power law model (10) and that given by an appropriate empirical relationship. In this project, the empirical relationship proposed by Witczak ([17]) is used:

$$log|E^*_{HMA}(f)| = -0.261 + 0.008225\rho_{200} - 0.00000101(\rho_{200})^2$$
$$+0.00196\rho_4 - 0.03157V_a - 0.415\frac{V_{beff}}{V_{beff}+V_a}$$
$$+\frac{1.87 + 0.002808\rho_4 + 0.0000404\rho_{38} - 0.0001786(\rho_{38})^2 + 0.0164\rho_{34}}{1 + e^{(-0.716logf - 0.7425log\eta)}} \quad (15)$$

where

$\|E^*_{HMA}\|$:	Amplitude of the complex Young's modulus of HMA, in 10^5 psi;
η:	Asphalt binder viscosity, in 10^6 poise (at any temperature, degree of aging);
f:	Frequency of the load, in Hz ($f = \frac{\omega}{2\pi}$, ω is the circular frequency);
V_a:	Percent air voids in the mix, by volume;
V_{beff}:	Percent effective bitumen content, by volume;
ρ_{34}:	Percent retained on $\frac{3}{4}$ in. sieve, by total aggregate weight (cumulative);
ρ_{38}:	Percent retained on $\frac{3}{8}$ in. sieve, by total aggregate weight (cumulative);
ρ_4:	Percent retained on No. 4 sieve, by total aggregate weight (cumulative);
ρ_{200}:	Percent passing No. 200 sieve, by total aggregate weight (cumulative).

The parameters that describe the HMA grading (V_a, V_{beff}, ρ_{34}, ρ_{38}, ρ_4, $\rho200$) are assumed to be known for a specific HMA and Table (4) lists their respective values for three characteristic types of HMA used by the Mn/DOT. Besides these parameters, the bitumen viscosity (η) in eq. 15 is calculated using the Superpave equation (Witczak, [17]) which is defined as:

$$\log_{10}(\log_{10}\eta_{T_{Rankine}}) = A + VTS(\log_{10}T_{Rankine}) \quad (16)$$
$$T_{Rankine} = T_{Fahrenheit} + 459.67$$

where $T_{Fahrenheit}$ is the asphalt temperature in Fahrenheit, A is the y-axis intercept (in the plot of $\log_{10}(\log_{10}\eta_{T_{Rankine}})$ against $(\log_{10}T_{Rankine})$) and VTS is the viscosity

temperature slope given by

$$VTS = \frac{\log\log\eta_{T_2} - \log\log\eta_{T_1}}{T_2 - T_1} \tag{17}$$

where T_1 and T_2 denote the binder temperature in Rankine. In general, parameters A and VTS depend on the type of HMA and Table (4) shows their values for three representative types of HMA.

Table 4: HMA Parameters

Type	A	VTS	$V_a[\%]$	$V_{beff}[\%]$	$\rho_{34}[\%]$	$\rho_{38}[\%]$	$\rho_4[\%]$	$\rho_{200}[\%]$
PG 58-28	8.1099	−2.6413	5.37	10.07	0	15.5	33.3	5.2
PG 58-34	8.8606	−2.9286	5.09	10.49	0	13.5	31.5	5.2
PG 58-40	9.2113	−3.0685	5.75	10.31	0	14	33	4.7

Based on the foregoing developments and a typical range of HMA grading parameters (see Table 5), the range of parameters D_0, D_1 and m in the power law model that are appropriate for simulating the viscoelastic behavior of asphalt concrete in Minnesota are listed in Table 6.

Table 5: Typical Ranges of HMA Parameters

Parameters	η_{binder} [10^6 Poise]	V_a [%]	V_{beff} [%]	ρ_{34} [%]	ρ_{38} [%]	ρ_4 [%]	ρ_{200} [%]
Minimum	0.00004	3.0	7.0	N/A	N/A	N/A	N/A
Medium	15000	6.0	11.0	N/A	N/A	N/A	N/A
Maximum	30000	9.0	15.0	N/A	N/A	N/A	N/A
Grade 1	N/A	N/A	N/A	0	10.0	20.0	7.0
Grade 2	N/A	N/A	N/A	5.0	50.0	60.0	5.0
Grade 3	N/A	N/A	N/A	45.0	70.0	85.0	2.0

Table 6: Range of Parameters in Power Law Model

Name	D_0 [10^{-11}Pa^{-1}]	D_1 [10^{-11}Pa^{-1}]	m
Minimum Value	4	2	0.15
Maximum Value	100	700	0.45

3.3 Aggregate Base, Subgrade and Half-space

For the aggregate base, subgrade and half-space soils, elastic Poisson's ratio (ν), mass density (ρ), and the hysteric (i.e. frequency-independent) damping ratios (ξ_s, ξ_p) do not vary greatly in typical field situations, and they are assumed to have constant values as shown in Table 7. The complex shear moduli (G^*) and complex Poisson's ratios (ν^*) for each layer that are required to simulate the FWD test using the dynamic model are calculated from the correspondence principle (Christensen, [18]) according to

$$G^* = (1 + 2\xi_s)G \tag{18}$$

$$\nu^* = \frac{2\xi_p(1-\nu) - 2\xi_s(1-2\nu) - i\nu}{4\xi_p(1-\nu) - 2\xi_s(1-2\nu) - i} \tag{19}$$

where elastic shear modulus (G) for each layer is specified in Table 7. As examined by

Table 7: Aggregate Base, Subgrade and Half-space Properties

Property	h [m]	G [MPa]	ν	ρ [kg/m^3]	ξ_s	ξ_p	E [MPa]
Aggregate Base	$0.1 \sim 1.0$	$10 \sim 130$	0.35	2027	0.01	0.005	$27 \sim 351$
Subgrade	$0.3 \sim 3.0$	$10 \sim 130$	0.4	1865	0.01	0.005	$28 \sim 364$
Half-space	N/A	$10 \sim 130$	0.45	2160	0.01	0.005	$29 \sim 377$

Ewing et.al. [14] among others, the hysteric damping model provides a close approximation of attenuation mechanisms in soils subjected to small-strain seismic waves (such as those generated in the FWD test).

4 Data Interpretation

4.1 Frequency Response Function (FRF)

4.1.1 Definition of FRF

To estimate the pavement stiffness properties from FWD measurements, the dynamic pavement response is characterized in the *frequency domain* using the concept of a frequency response function (FRF) whose values are taken as input to the back-analysis. If the temporal variation of an input to a linear system is denoted as $x(t)$ and its output is denoted as $y(t)$, the frequency response function characterizing the system is defined as:

$$FRF(f) = \frac{Y(f)}{X(f)} \tag{20}$$

where $Y(f)$ and $X(f)$ are the Fourier transforms of $y(t)$ and $x(t)$, respectively (see Fig. 6). In the above, the forward and inverse Fourier transform of a given time history $h(t)$ are defined as

$$\begin{aligned} H(f) &= \int_{-\infty}^{\infty} h(t) e^{-i2\pi ft} dt \\ h(t) &= \int_{-\infty}^{\infty} H(f) e^{i2\pi ft} df \end{aligned} \tag{21}$$

where $i = \sqrt{-1}$ (see [15]).

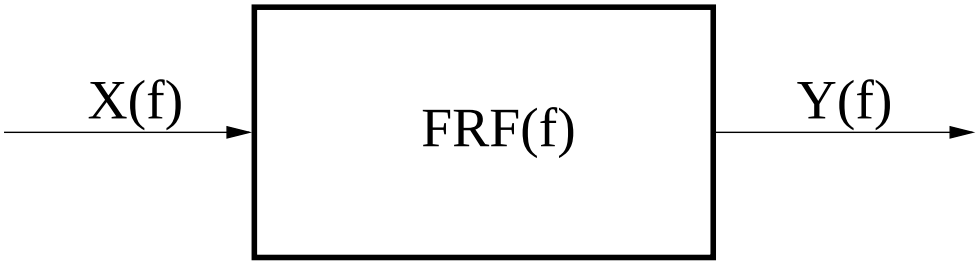

Figure 6: Linear System

4.1.2 Selection of Input and Output Signals

As demonstrated in Guzina and Nintcheu [11], the contact pressure distribution between the pavement and the loading plate in an FWD test is typically non-uniform. This interaction effect, which is in contradiction with the assumption of the forward model employed, is known to affect the response of the center geophone (G_0) significantly. As a result, the deflection measured at the center geophone is eliminated from consideration in this study. Moreover, due to the inertia of the loading plate (see Fig. 7), the force $F(t)$ measured on top of the loading plate is not equal to the actual force $Q(t)$ acting on the pavement, and thus may not represent a reliable measure of dynamic excitation of the pavement system. To alleviate the problem, instead of taking the load signal as an input and deflection signal as the output, the displacement time history at geophone p (denoted as $d_p(t)$) is taken as the input signal while the displacement time history at geophone q (denoted as $d_q(t)$) is taken as the output. As a result, the transfer functions characterizing the pavement system in this study can be written in compact form as

$$FRF_{pq} = \frac{D_q(f)}{D_p(f)} \quad , \qquad q > p \tag{22}$$

where $D_q(f)$ and $D_p(f)$ are the Fourier transforms of $d_q(t)$ and $d_p(t)$, respectively. In this project, the pair (p, q) takes two sets of values, namely $(p = 2, q = 5)$ and $(p = 5, q = 8)$ with reference to geophones 2, 5 and 8 in Fig. 2.

4.1.3 Usable Frequency Range

Due to the finite width of the pavement section, it is reasonable to expect that lateral reflections from the edges of the asphalt concrete layer would influence the FWD measurements. Unfortunately, such reflections could not be accounted for by the forward model employed which assumes that all pavement layers infinitely extend in the lateral direction. In the frequency domain, this effect takes the form of an unusual peak around 50 to 60 Hz, and Fig 9 shows one such example.

To substantiate the foregoing argument, the fundamental resonant frequency for transverse vibration of the HMA layer of width w_{HMA} can be estimated by approxi-

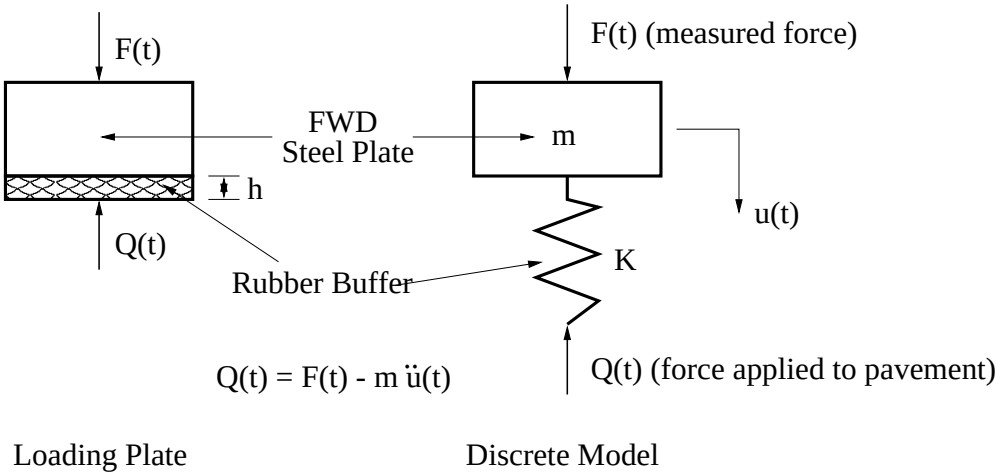

Figure 7: Illustration of Inertial Effects in FWD Testing

mating the asphalt layer as a free-free rod as shown in Fig. 8. According to the one-dimensional theory of wave propagation, the kth resonant frequency of the free-free rod in transverse vibration can be written as

$$f_k = \frac{k \; c_{s,HMA}}{2w_{HMA}} \tag{23}$$

where k is the mode number, $c_{s,HMA}$ is the shear wave speed of asphalt concrete, and w_{HMA} is the width of the pavement. By letting $k = 1$, the fundamental resonant frequency is

$$f_{resonant} = \frac{c_{s,HMA}}{2w_{HMA}} \tag{24}$$

As a result, for the test section examined in Fig 9, the fundamental resonant frequency of the asphalt concrete layer in transverse vibration can be approximated by

$$\begin{aligned} f_{resonant} &= \frac{c_{s,HMA}}{2w_{HMA}} \\ c_{s,HMA} &= \sqrt{\frac{G_{HMA}}{\rho_{HMA}}} \\ G_{HMA} &= \frac{E_{HMA}}{2(1+\nu_{HMA})} \end{aligned} \tag{25}$$

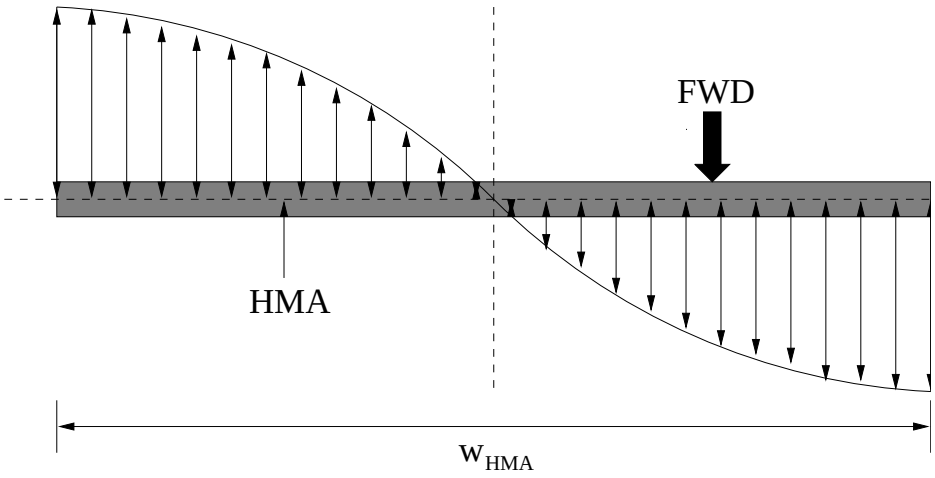

Figure 8: Fundamental Vibration Mode for the Free-Free Rod

where $c_{s,HMA}$ is the shear wave speed of asphalt concrete, w_{HMA} is the width of the pavement, which in this case is 7.32 meters (24 feet), G_{HMA} is the elastic shear modulus of HMA, $\nu_{HMA} = 0.4$ is the Poisson's ratio of HMA, $\rho_{HMA} = 2335$ kg/m^3 is the mass density of the HMA, and E_{HMA} is the elastic Young's modulus of HMA with an assumed value of 5 GPa. Based on these assumptions, the fundamental resonant frequency of the asphalt layer can be can be calculated to be $f_{resonant} = 60$ Hz.

To alleviate the problem associated with lateral reflections, the frequency range used in this study is taken as:

$$[f_{min}, f_{max}] = [1, 51] \text{ Hz} \tag{26}$$

4.1.4 Reliability of the FRF Measurements

The FWD device used by the Mn/DOT (i.e. the loading plate, falling weight, geophones and other components) is built on a steel frame (see Fig. 11). To check whether the vibration transmitted through the frame has any effect on the deflection measurements, a comparison test was made by comparing the frequency response functions obtained using the frame-mounted FWD geophones and the FRF obtained using the high fidelity

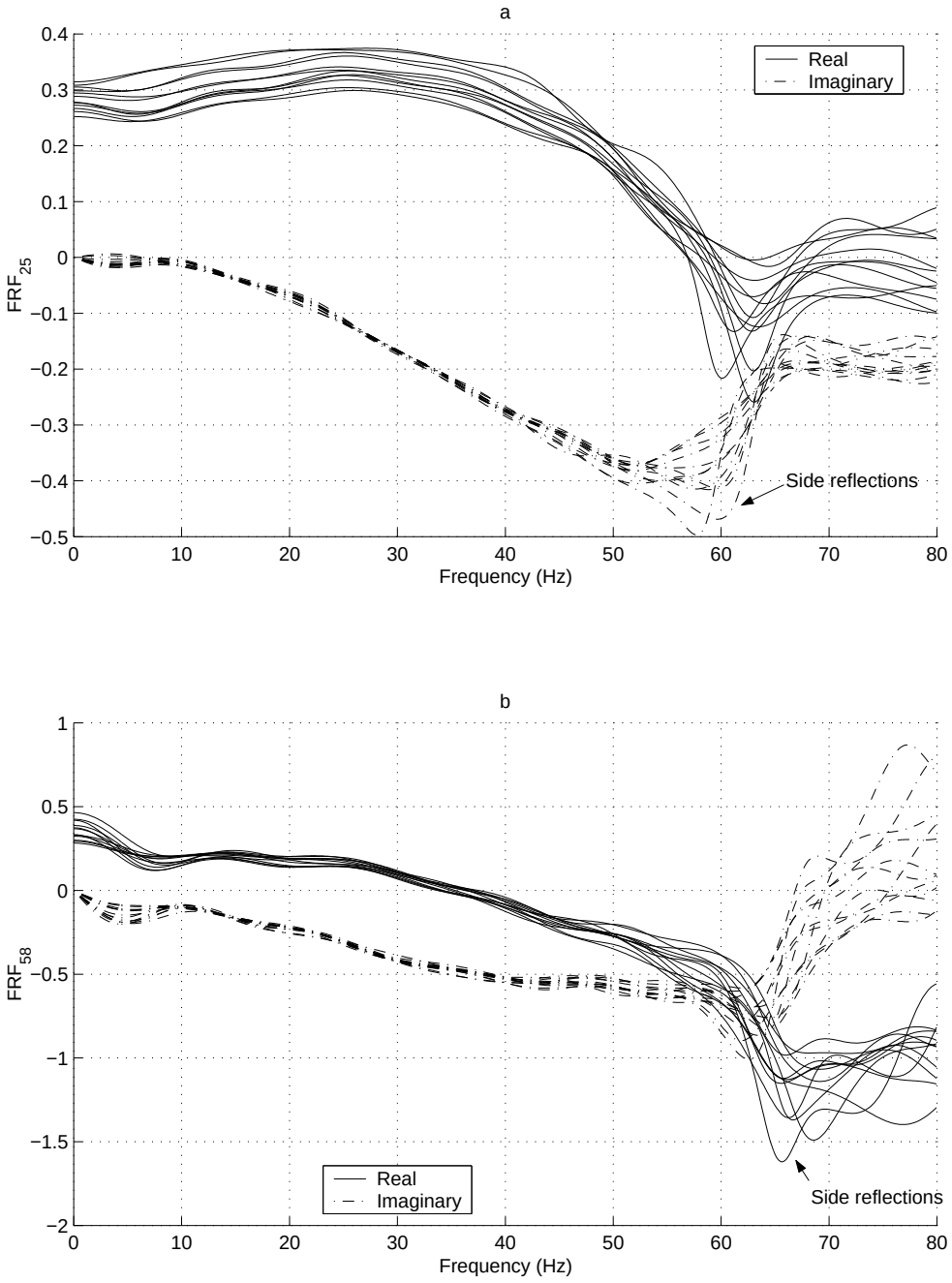

Figure 9: FRF in FWD Test (Test Section 33, May 22, 2001)

accelerometers mounted independently at the same position. In the comparison, the FWD geophones were used first and the frequency response functions were obtained from the deflection time histories as measured by the geophones. Second, the geophones

2, 5 and 8 are lifted and three accelerometers were glued onto the pavement at the same place corresponding to the three geophones. Third, the FWD test with the same load level was performed and the frequency response functions were (22) re-computed from the acceleration time histories. The comparison is shown in Fig. 10. From the Figure, it can be seen that:

- The common feature of the FRF's measured using both geophones and accelerometers is that there is an unusual resonance peak located around 50 Hz to 60 Hz. This means that this peak is not induced by the FWD device, but rather by the pavement characteristics. Therefore, the above speculation that this peak is due to the finite width of the pavement appears to be reasonable. Also, the device that was used to record the vibrations with the aid of high fidelity accelerometers is DC powered, which eliminate the possible influence of AC electrical noise.

- It should be noted that the frequency response functions given by the two sets of transducers differ significantly in the low frequency range (below 10 Hz). The reason for this irregularity is that the accelerometers are not well-suited for vibration measurements at very low frequencies (Nazarian and Stokoe, [19]) owing to the high noise-to-signal ratio. One may also observe that this irregularity is more pronounced for $FRF_{8/5}$ than $FRF_{5/2}$, which can be explained by the lower signal-to-noise ratio at the location of geophone 8 than that at geophone 5.

4.1.5 FRF Sampling Scheme

To use the selected frequency response functions as an input to the backcalculation analysis, the simplest way is to take the values of the FRF of interest at several discrete frequencies as an input. This is shown in Fig. 12a, where the real and imaginary components of the FRF at five evenly distributed frequency points are taken as an input. However, this strategy may not be robust when dealing with actual data because of (i) the inevitable presence of noise in the FWD measurements, and (ii) the fact that frequency response function represented by the five sampling frequencies would be

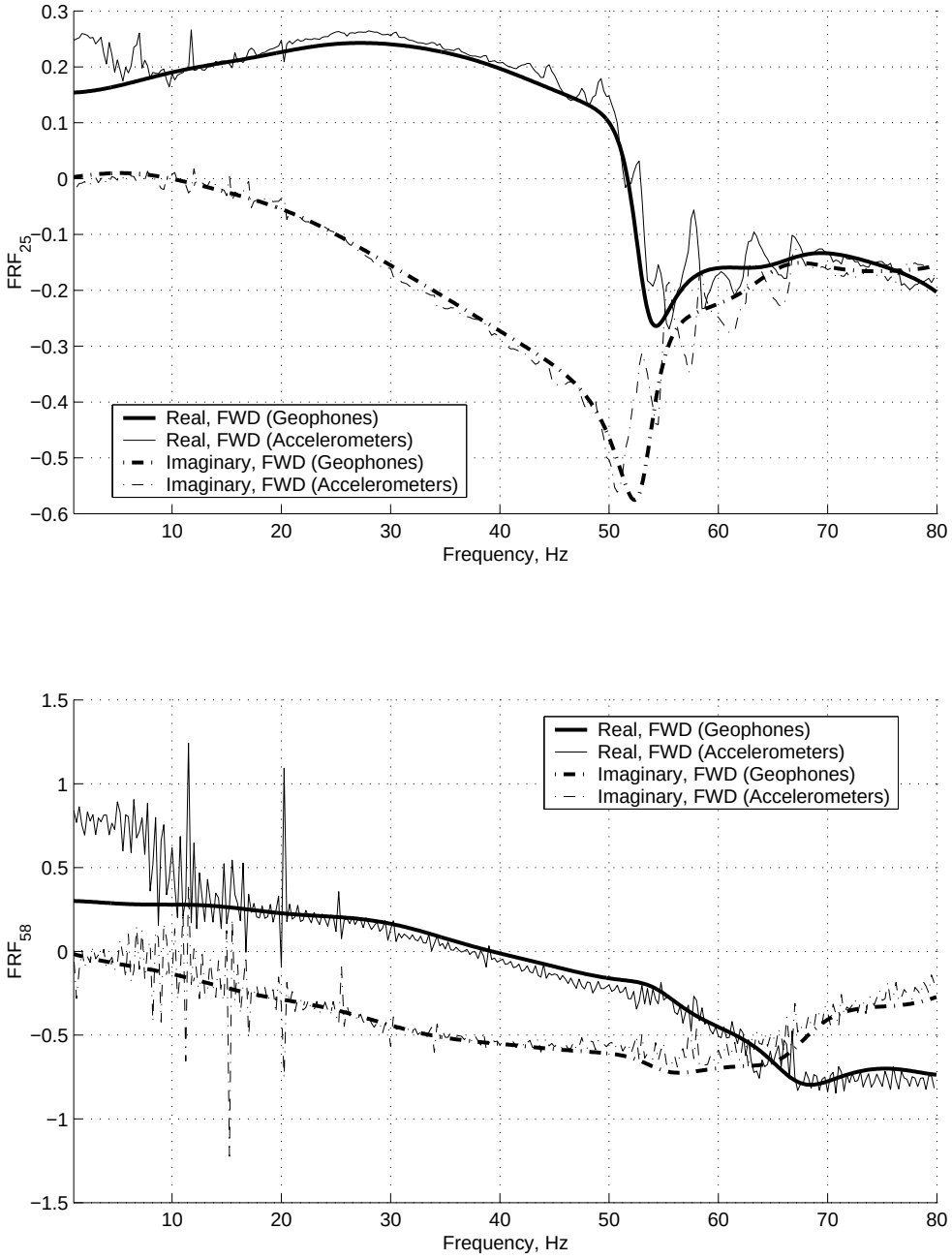

Figure 10: Comparison of FWD using Geophones and Accelerometers

influenced greatly even if there is only a slight level of noise around the sampling frequencies. An alternative sampling scheme is illustrated in Fig. 12b. In this approach, the frequency range of interest is divided into several segments. For each segment, the

Figure 11: Dynatest FWD Device used by the Minnesota Department of Transportation

areas under the real and imaginary components of the FRF over that segment are taken as an input to the backcalculation analysis. This strategy is more robust than the point sampling scheme because the frequency response function represented by the segmental areas will not be as sensitive to the noise at several isolated frequency points.

4.2 FRF from Forward Model

The visco-elastodynamic forward model employed in this study simulates the FWD test in the frequency domain. For a given set of pavement properties at a certain frequency, the FRF's of interest at this specific frequency can be calculated. To obtain the variation of FRF's over a certain frequency range, the FRF must be evaluated multiple times within this range. However, due to the sampling scheme adopted in the current project (see Section 4.1.5), it is important that the simulated frequency response functions be evaluated using a sufficient density of sampling points so that the integration results have an acceptable level of error. However, it will be impractical to simulate the FWD test

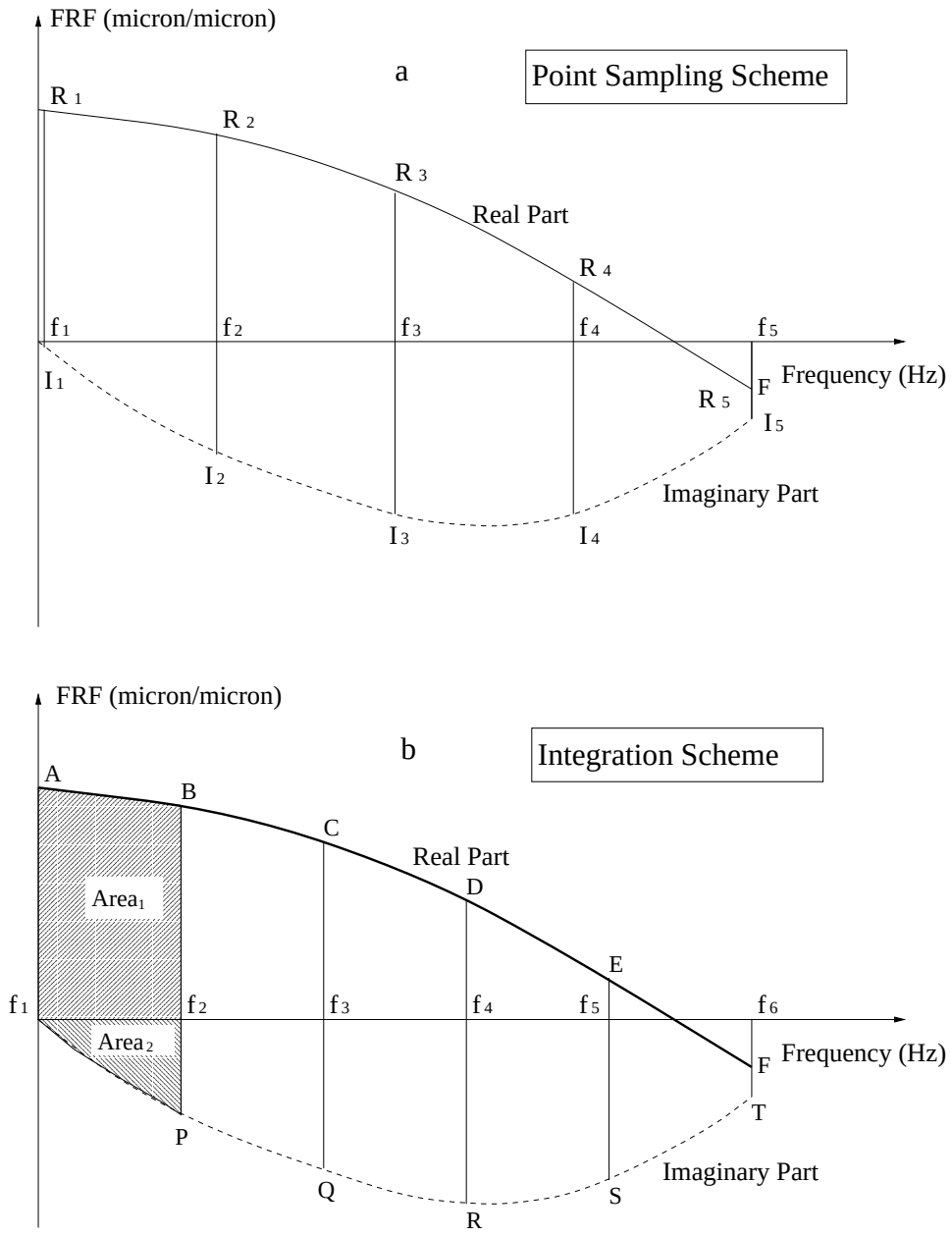

Figure 12: Integration Scheme

using a very large number of frequencies. To determine the optimal number of sampling frequencies for the theoretical, i.e. forward model, a sensitivity study is performed as shown in Fig. 13. In this comparison, the *"exact"* integration was based on the frequency response function represented by 2000 frequencies, which is a far greater number than

needed. In the figure, the integration error is defined as:

$$(Error_{integration})_{N_j} = \frac{1}{N_{area}} \sum_{i=1}^{N_{area}} \left| \frac{Area_{i,2000} - Area_{i,N_j}}{Area_{i,2000}} \right| \tag{27}$$

where N_{area} is the number of integration intervals times 2 (for real and imaginary parts), $Area_{i,N_j}$ is the ith area under the frequency response function calculated when N_j frequencies are used, $Area_{i,2000}$ is the ith area under the frequency response function calculated when 2000 sampling frequencies are used, and $(Error_{integration})_{N_j}$ is the integration error when N_j frequencies are used in calculating the area under the frequency response function. From Fig. 13, it can be seen that $N_j = 50$ is sufficiently large to minimize the integration error.

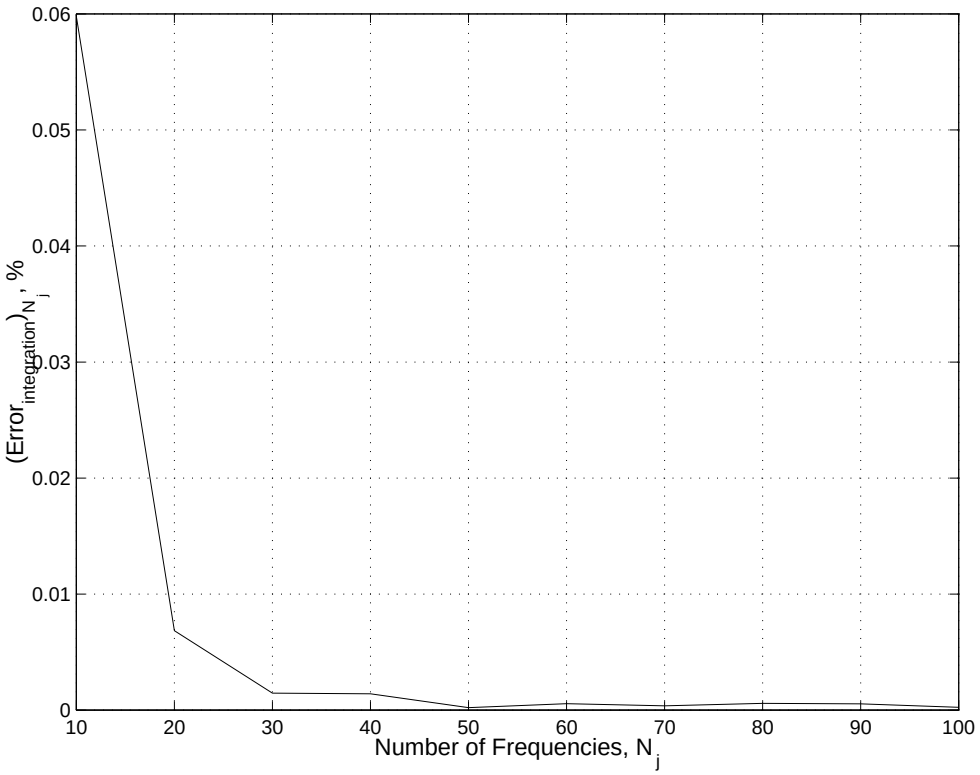

Figure 13: Selection of the Number of Sampling Points Needed for FRF Integration

4.3 FRF from FWD measurements

Because the experimental (i.e. field) frequency response functions are used as an input to the backcalculation scheme developed in this project (as they encode the pavement stiffness and damping properties), it is imperative that the FRF's be computed from FWD measurements effectively and accurately.

4.3.1 FRF Estimates

In general, there are two basic ways to estimate the frequency response function of interest from the measured deflection time histories.

- **Direct Division:** This method is directly based on eq. (22). Because the recorded time history is a digitized signal which includes only values at discrete points, the discrete Fourier transform (DFT) is used to calculate $D_p(f)$ and $D_q(f)$ featured in (22). For test j ($j = 1, 2, ..., n$, where n is the total number of FWD tests), the DFT of the displacement time histories at geophones p and q is calculated respectively as

$$[D_p(f_k)]_j = \Delta t \sum_{m=0}^{M-1} [d_p(t_m)]_j e^{-i\frac{2\pi km}{M}} \quad , \quad k = 0, 1, 2, ..., M-1 \tag{28}$$

$$[D_q(f_k)]_j = \Delta t \sum_{m=0}^{M-1} [d_q(t_m)]_j e^{-i\frac{2\pi km}{M}} \quad , \quad k = 0, 1, 2, ..., M-1$$

  where Δt is the sampling interval, M is the length of the discrete Fourier transform record, and $f_k = \frac{k}{M\Delta t}$ (see also Bendat and Piersol, [20]). On the basis of (28),the averaged FRF can be written as:

$$FRF_{pq}(f) = \frac{1}{n} \sum_{i=1}^{n} \frac{[D_q(f_k)]_i}{[D_p(f_k)]_i} \quad , \quad q > p \tag{29}$$

- **Spectral Average:** In this method, the FRF is estimated using the concepts of the cross spectral density and the auto spectral density (Bendat and Piersol, [20]) in the form of

$$FRF_{pq} = \frac{G_{pq}(f)}{G_{pp}(f)} \tag{30}$$

where $G_{pq}(f)$ is the cross spectral density between the displacement signal at geophone p and the one at geophone q, while $G_{pp}(f)$ is the auto-spectral density of the displacement signal at geophone p. In practice, $G_{pq}(f)$ and $G_{pp}(f)$ are estimated via (Bendat & Piersol, [20])

$$G_{pq}(f_k) = \frac{1}{n}\sum_{j=1}^{n}[D_p^*(f_k)]_j[D_q(f_k)]_j \tag{31}$$

$$G_{pp}(f_k) = \frac{1}{n}\sum_{j=1}^{n}[D_p^*(f_k)]_j[D_p(f_k)]_j \tag{32}$$

where n is the total number of tests, $i = 1, 2, ..., n$ is the test number, $D_p(f)$ is the Fourier transform of the deflection signal at geophone p, and $D_q^*(f)$ is the complex conjugate of $D_q(f)$. To eliminate the effect of side leakage in eq. (28), $[D_p(f)]_i$ is calculated via

$$[D_p(f_k)]_j = \Delta t \sum_{m=0}^{M-1} v(t_m)[d_p(t_m)]_j e^{-i\frac{2\pi km}{M}} \quad k = 0, 1, 2, ..., M-1 \tag{33}$$

where $[d_p(t)]_i$ is the signal at geophone p for test i, $v(t)$ is a window function used to smooth the estimation, and C is the loss factor caused by a given window function $v(t)$. In general, C can be defined as

$$C = \frac{\int_0^T v^2(t)dt}{\int_0^T v_{box}^2(t)dt} \tag{34}$$

where T is the signal duration of, and $v_{box}(t)$ is the boxcar window function. A definition analogous to (33) applies to $[D_q(f)]_i$. In (34), all window functions are assumed to be zero for $t < 0$ and $t > T$. For $0 \le t \le T$, the most commonly used window functions are defined as

$$\begin{aligned}
v_{box}(t) &= 1 && && \text{Boxcar Window} \\
v(t) &= 0.5(1 - cos\frac{2\pi t}{T}) && && \text{Hanning Window} \\
v(t) &= \frac{2t}{T} && 0 \le t \le \frac{T}{2} && \text{Triangular Window} \\
v(t) &= \frac{2(T-t)}{T} && \frac{T}{2} \le t \le T &&
\end{aligned} \tag{35}$$

Table 8: Pavement Properties Used in Windowing Function Comparison

Property	h [m]	G [MPa]	ν	ρ [kg/m^3]	ξ_s	ξ_p	E [MPa]
HMA	0.1	1200	0.35	2335	0.01	0	3240
Aggregate Base	0.3	80	0.35	2027	0.005	0	216
Subgrade	1.0	40	0.4	1865	0.005	0	112
Half-space	N/A	120	0.45	2160	0.005	0	348

Together with the direct division method, the spectral average method with three window functions (Boxcar, Hanning and Triangular) is used to evaluate the FRF's from a synthetic displacement time history record for which the exact FRF's are known. The synthetic displacement time history is generated using the forward model in Guzina & Pak [4] based on profile shown in Fig. 5 with pavement properties listed in Table 8. From Fig. 14, which plots the frequency response functions obtained via the respective methods, it can be concluded that the boxcar window is the best choice for the current project.

4.3.2 Quality of FWD Measurements

In order to interpret the field data and backcalculate the pavement properties reliably, it is necessary to ensure that the field measurements have an acceptable accuracy. The coherence function is an effective criterion that can be used for this purpose. Upon assuming the deflection time history at geophone p to be an input signal and the deflection time history at geophone q to be an output signal, the coherence function is defined as (Bendat & Piersol, [20]):

$$\gamma^2(f) = \frac{|G_{pq}(f)|^2}{G_{pp}(f)G_{qq}(f)} \tag{36}$$

where G_{pp} denotes the auto-spectral density, and G_{pq} is the cross-spectral density function as given in Section 4.3.1.

For a linear system with no extraneous noise, $\gamma^2(f)$ is unity for all frequencies. If the system is non-linear or noise is present in the measurements, $\gamma^2(f)$ is smaller than

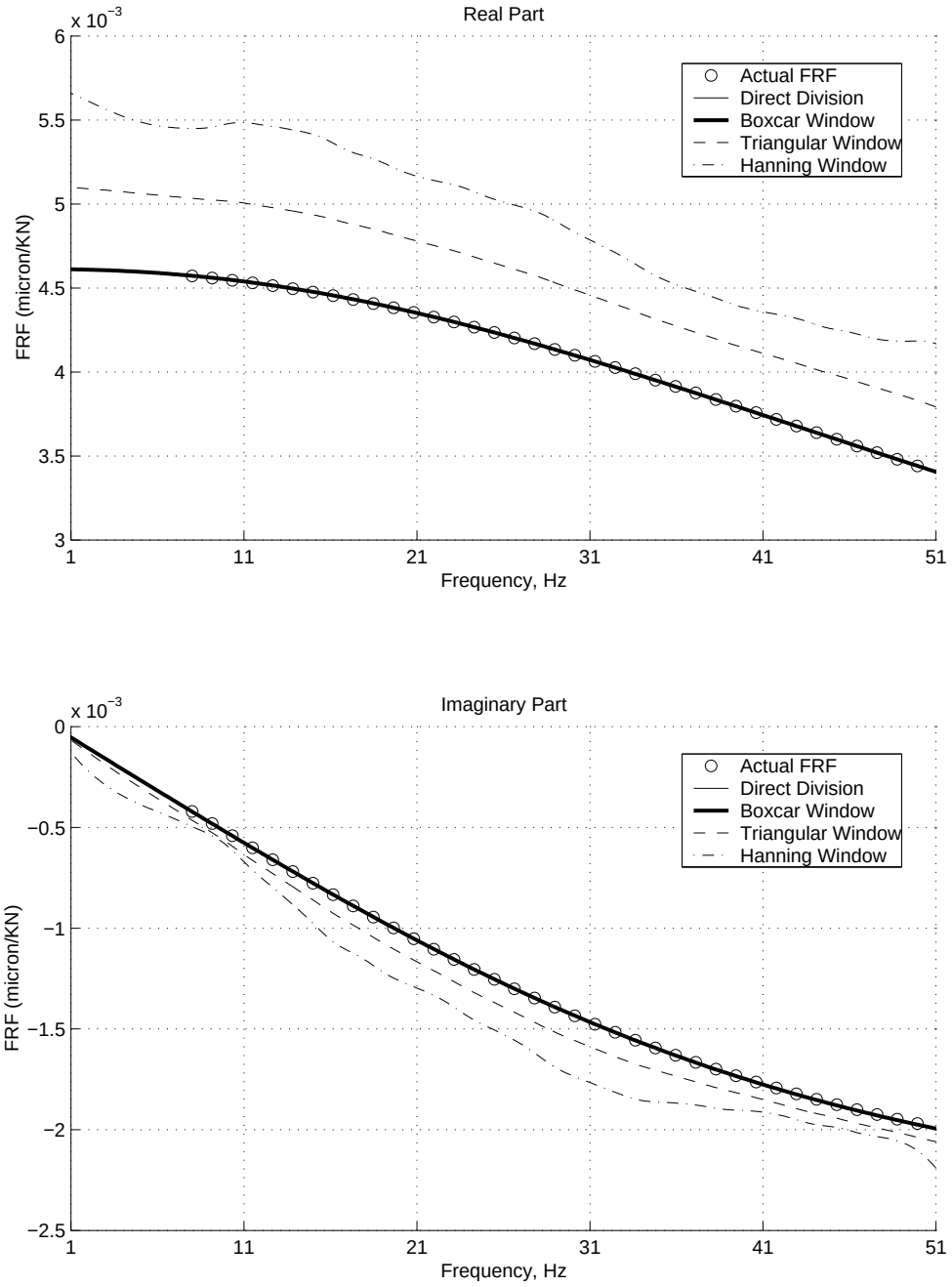

Figure 14: Comparison of Window Functions

unity. Therefore, the measured data should be considered acceptable only if $\gamma^2(f)$ is equal or close to unity. Figs. 15 and 16 illustrate this criterion. In Fig. 15, the coherence function is significantly smaller than unity (especially at lower frequencies)

and this measurement should be discarded. In Fig. 16, the coherence is close to unity and the associated measurements can be considered to have an acceptable level of noise. The foregoing method of input data quality checking has been incorporated into the backcalculation software developed in this study.

It should be noted, however, that the coherence criterion is only a measure of the relative noise in experimental observations, and not a check of the measurement accuracy. Based on the study of the field FWD measurements, another criterion for estimating the quality of the FWD measurements has been established that is based on the deflection time history measured by the center geophone. When this deflection time history oscillates below and above zero after the main peak (Fig. 17a), this measurement is judged to be acceptable. Otherwise, if it hovers above zero after the main peak (Fig. 17b), this measurement is believed to have a poor quality.

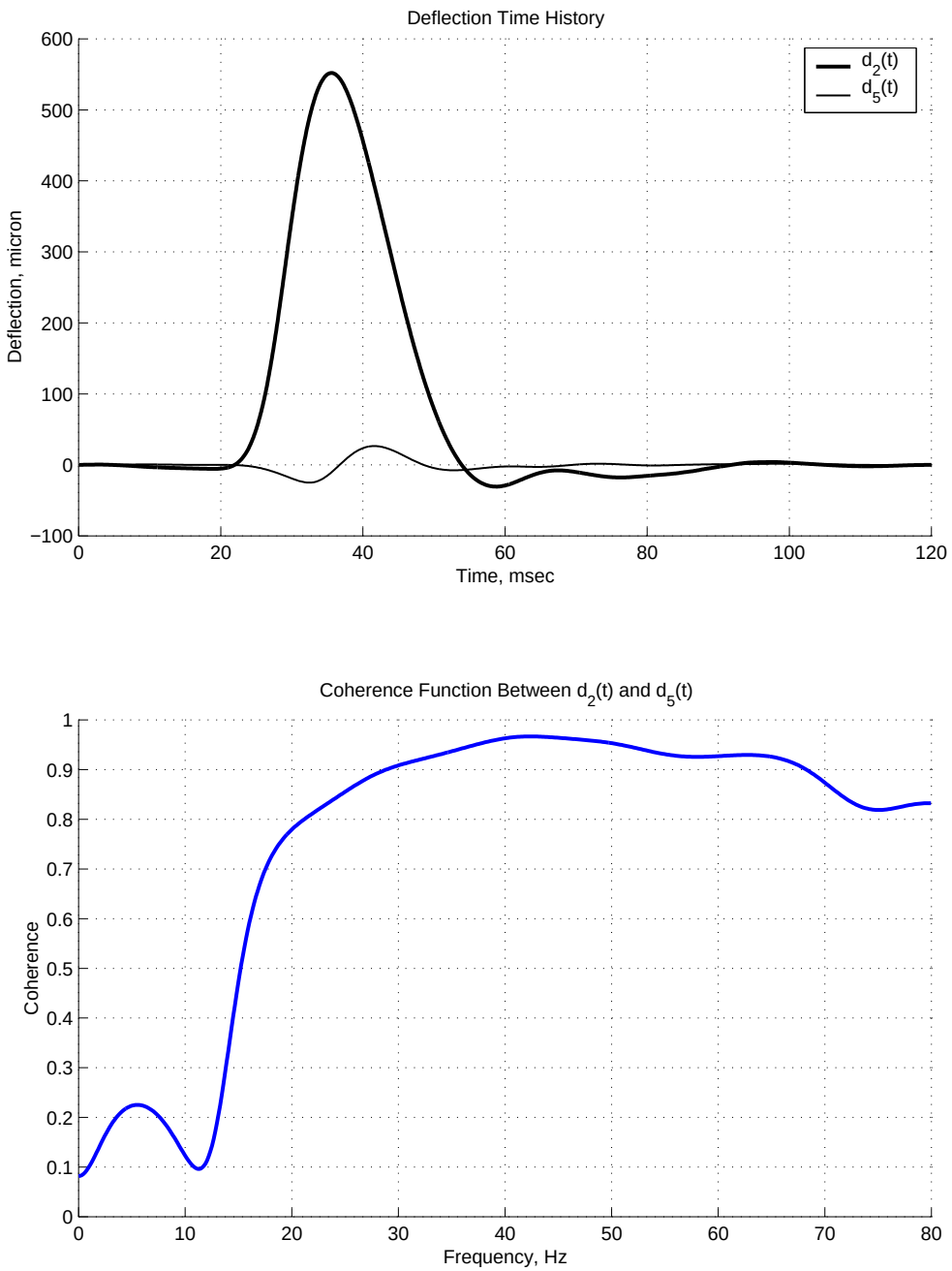

Figure 15: Coherence Illustration for Test Section 28, Aug. 4th, 1999

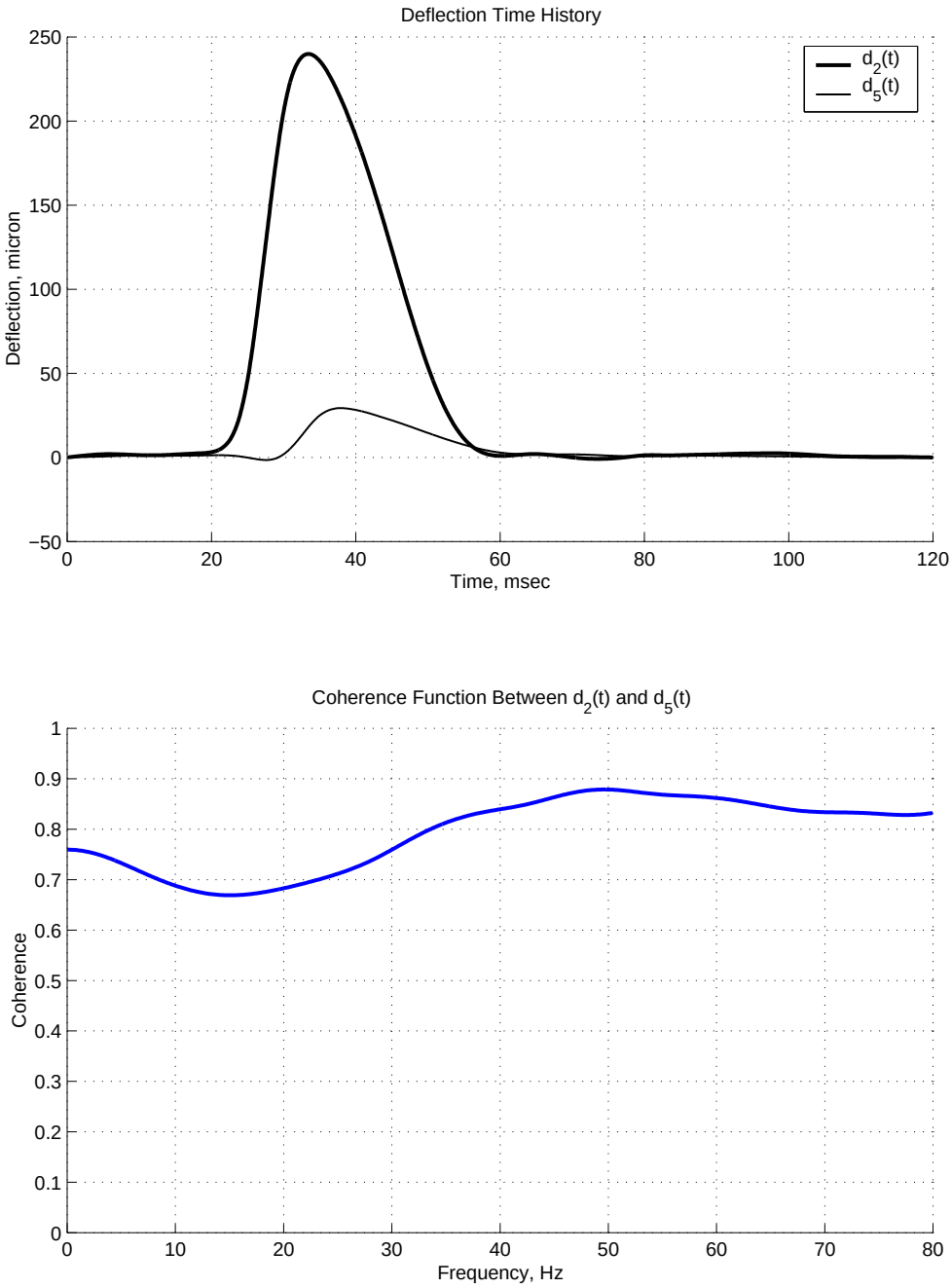

Figure 16: Coherence Illustration for Test Section 35, Aug. 4th, 1999

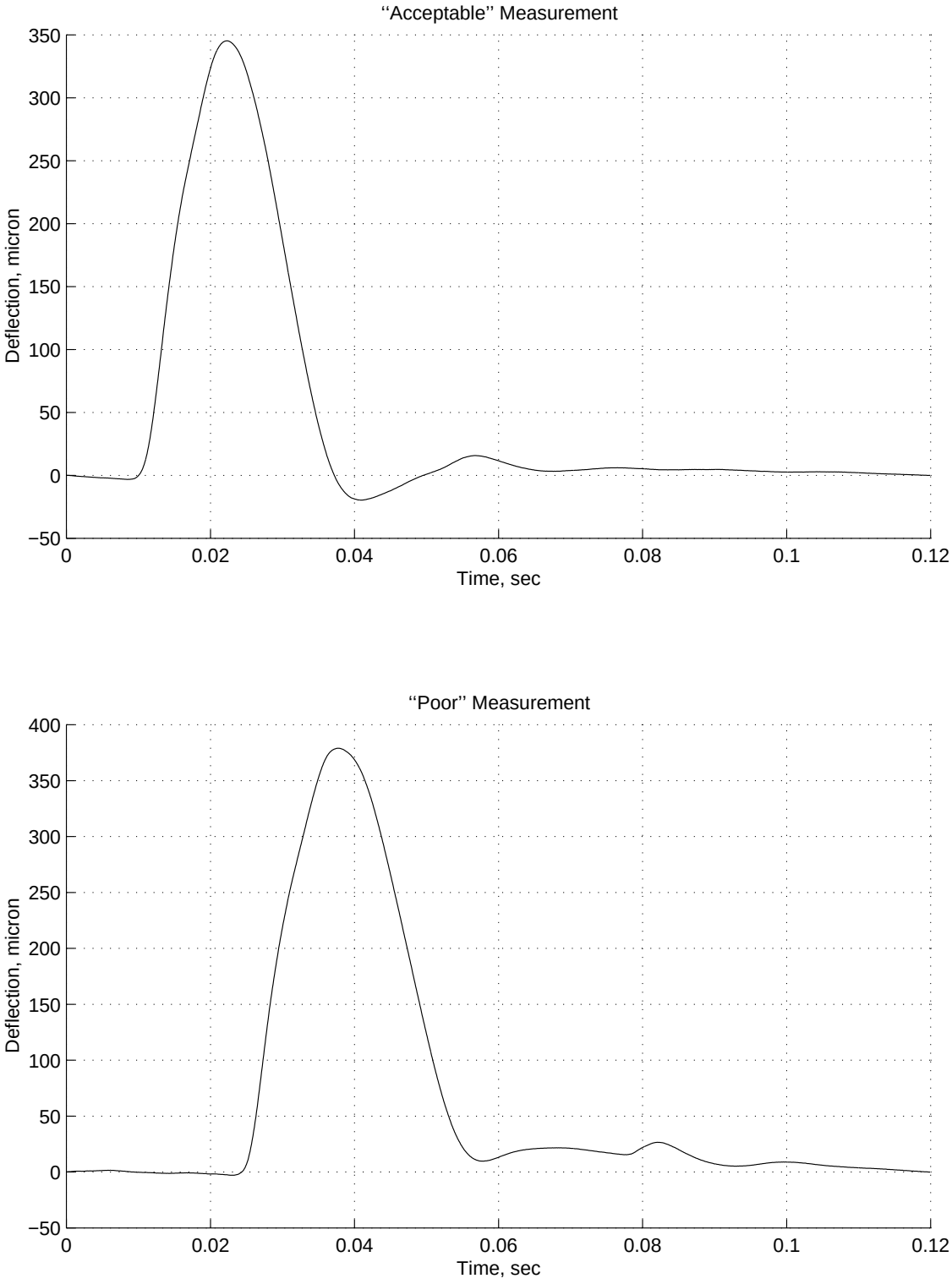

Figure 17: Deflection Record at Center Geophone Used as a Criterion for Judging the Measurement Quality

4.3.3 Baseline Correction

As stated above, the FRF's (FRF_{25} and FRF_{58}) are estimated from the deflection time histories at selected geophones (G_2, G_5 and G_8) as measured by the FWD. Because the FWD apparatus actually measures the particle *velocity* at geophone locations, the velocity time histories must be integrated to obtain the pavement deflections. However, owing to the noise embedded in the measured velocity time histories, such integration will cause the measurement error to be accumulated. As shown in Fig. 18, the accumulated noise results in a deflection time history that has a non-zero value at the end, which is not reasonable for a linear pavement system. The purpose of baseline correction is to delete this accumulated noise from the deflection time history and three such approaches are shown below. In these three methods, the deflection time history starts at time t_0 and ends at time t_1, and the corresponding deflections $d(t)_{t=t_0}$ and $d(t)_{t=t_1}$ are denoted as d_0 and d_1.

- **Linear Correction** The noise $\delta(t)$ embedded in a deflection time history $d(t)$ is assumed to take the linear temporal variation, and the deflection time history after baseline correction ($d'(t)$) can be calculated as:

$$d'(t) = d(t) - \frac{d_1 - d_0}{t_1 - t_0} t \tag{37}$$

- **Quadratic Correction** The noise $\delta(t)$ embedded in a deflection time history $d(t)$ is assumed to take the quadratic temporal variation, and the deflection time history after baseline correction ($d'(t)$) can be calculated as:

$$d'(t) = d(t) - \frac{d_1 - d_0}{(t_1 - t_0)^2} t^2 \tag{38}$$

- **Square Root Correction** The noise $\delta(t)$ embedded in a deflection time history $d(t)$ is assumed to take the square root temporal variation, and the deflection time history after baseline correction ($d'(t)$) can be calculated as:

$$d'(t) = d(t) - \frac{d_1 - d_0}{\sqrt{t_1 - t_0}} \sqrt{t} \tag{39}$$

Fig. 19 shows the effects of baseline correction on the estimated frequency response functions. From the figure, it can be seen that (i) the baseline correction greatly improves the estimated FRF, and (ii) FRF's based on three baseline correction methods do not differ significantly at frequencies above 10 Hz. Fig. 20 shows the effects of baseline correction on the coherence function. From the Figure, the beneficial effect of baseline correction should again be noted.

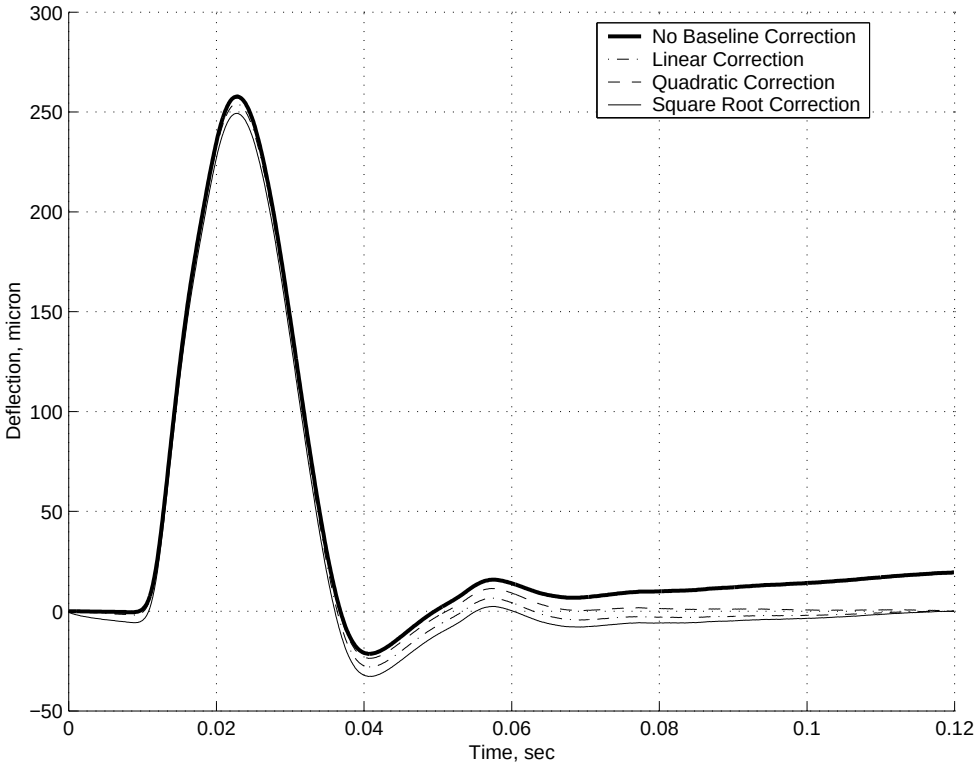

Figure 18: Effects of Baseline Correction On Deflection Time History (Test Section 33, 05/22/2001)

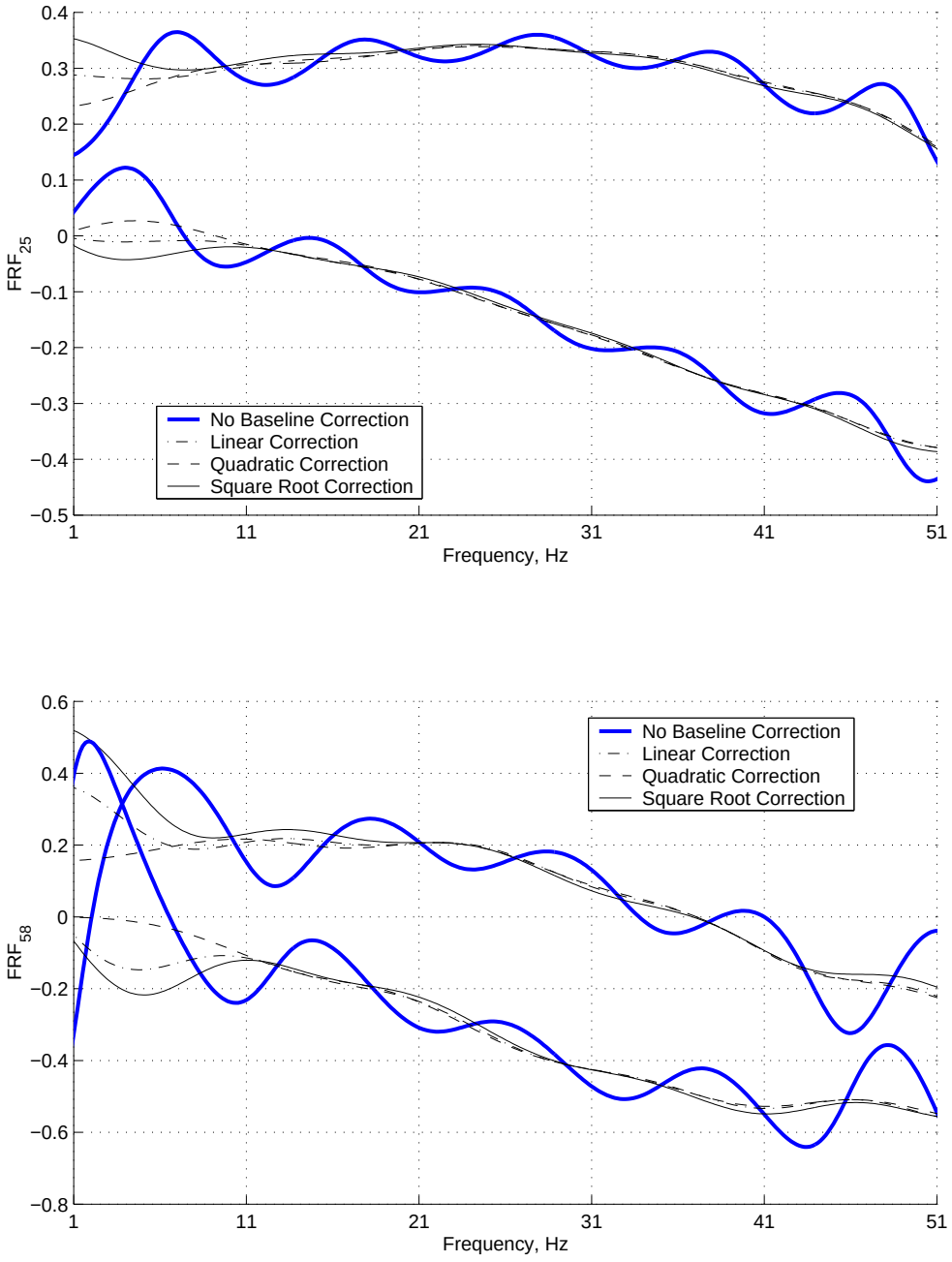

Figure 19: Effects of Baseline Correction on the Frequency Response Functions

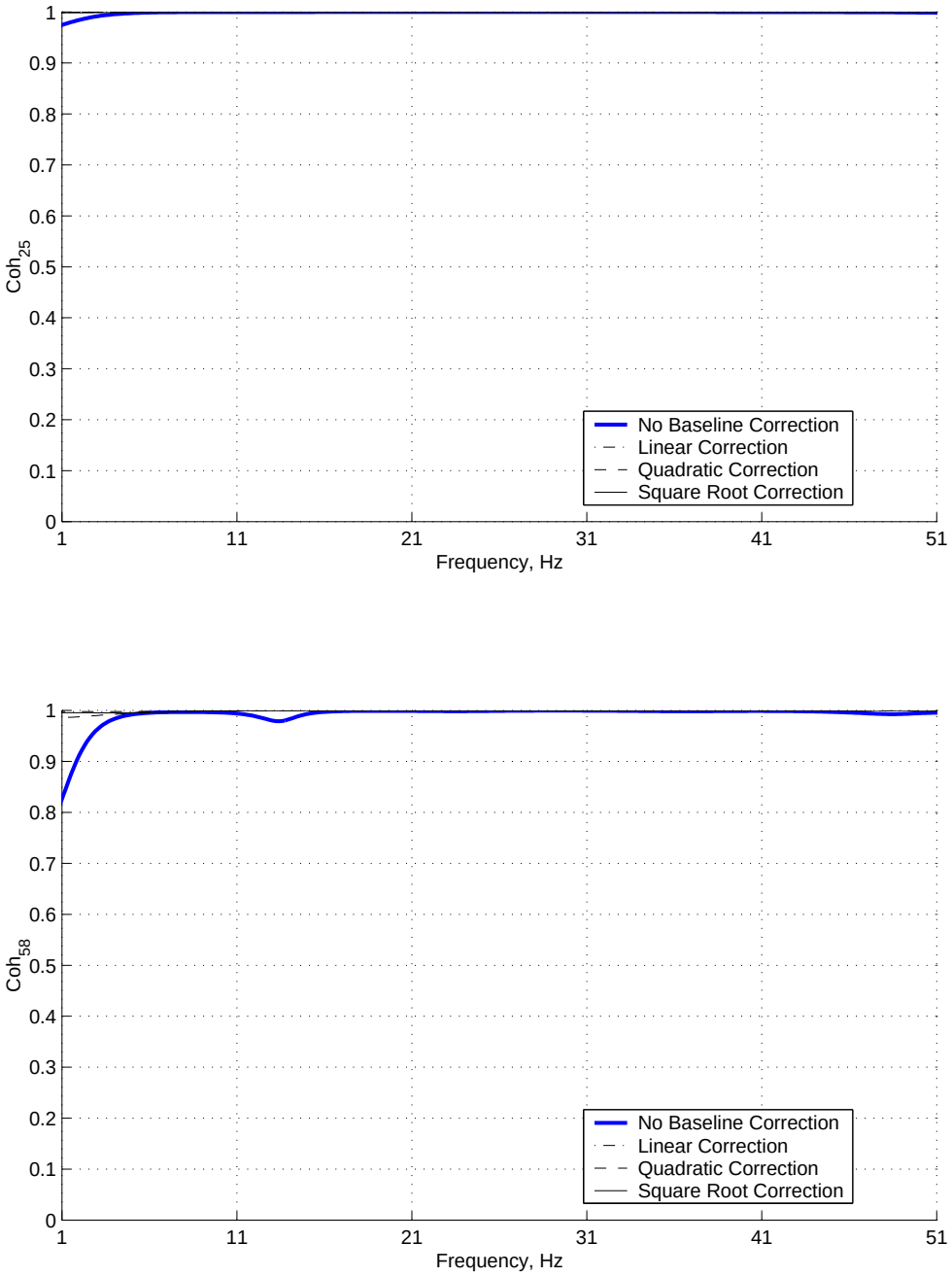

Figure 20: Effects of Baseline Correction on the Coherence Function

5 Artificial Neural Network (ANN)

5.1 Introduction to ANN

Artificial Neural Network (ANN) is a relatively new computational tool that is well suited for an effective treatment of a variety of inverse problems. With reference to Fig. 21 showing the architecture of a commonly used three-layer, feed-forward neural network, the ANN can be described as follows:

1. A neural network consists of many interconnected processing elements called *neurons*. For the multilayer feed-forward configuration used in this investigation, the ANN takes a special form. It has several layers, and each layer has a number of neurons. Every neuron is connected to the neurons of the neighboring layers only. The connection is one-direction, which means that the information is passed only from layer $k-1$ to layer k, and not vice versa. There are also no connections among neurons in the same layer. The strength of connection between the two neurons is represented by the *weight* attached to this connection. For the connection from neuron j to neuron i, the corresponding weight is defined as $w_{i,j}$.

2. Each neuron performs a simple computation on incoming signals in three steps (Russell and Norvig, [21]). First, it calculates the net input using (here neuron s in layer k is used as an example)

$$n_s^k = \sum_{m=1}^{S_{k-1}} w_{s,m}^k a_m^{k-1} + b_s^k \tag{40}$$

where n_s^k is the net input to neuron s; S_{k-1} is the number of neurons in layer $k-1$; $w_{s,m}^k$ is the weight from neuron m to neuron s; a_m^{k-1} is the output of neuron m in layer $k-1$, and b_s^k is the bias of neuron s in layer k. Secondly, the neuron calculates its own output (a_s^k) by applying the activation function of layer k to the net input (n_s^k) via

$$a_s^k = f^k(n_s^k) \tag{41}$$

Theoretically, activation function $f^k(x)$ can take any form, such as the threshold function and Gaussian function, and each neuron can have a different activation function. In the current project, the neurons in the last layer (i.e. output layer) use the linear activation function [22] given by

$$f^k(x) = x \tag{42}$$

while the neurons in all other layers (excluding input layer) use the hyperbolic tangent sigmoid function defined as

$$f^k(x) = \frac{2}{1 + e^{-2x}} - 1, \qquad k = 1, 2 \tag{43}$$

Third, this output is propagated to the neurons in the next layer $(k + 1)$ through its output connections, which in turn becomes the incoming signal for the neurons in the next layer.

3. Neural network receives its inputs at the input layer. The number of neurons in the input layer is equal to the number of input parameters for a specific inverse problem. The neurons of input layer forward the signals directly to the next layer without performing any other operations. The input layer is typically labeled as layer 0.

4. For each layer in the network except the input one, there is a special neuron called bias neuron with a fixed output of unity. The bias neuron is connected to every neuron in its layer, and the weight of these connections is adjusted during ANN training. For layer k, the weight from the bias neuron to neuron s is labeled as b_s^k. The use of a bias neuron is not required for some problems, but it is generally beneficial in most cases. This can be illustrated as follows (see also [23]). For a fully connected feed-forward neural network, a neuron in layer k will have S_{k-1} inputs (there are S_k neurons in layer k and S_{k-1} neurons in the layer $k - 1$), which defines a S_{k-1}-dimensional space. This neuron draws a hyperplane through the corresponding S_{k-1}-dimensional space, whose position is determined by the

weights from the previous layer to this neuron and the weight from the bias neuron in layer k. For the entire layer k, there are S_k such hyperplanes. Without the use of a bias neuron, all the hyperplanes are constrained to pass through the origin. This constraint is appropriate for some cases, but not appropriate for many problems such as the one shown in Fig. 22. In Fig. 22, a neural network is trained to classify the points labeled with "+" from the points labeled with "o". An appropriate line that separates these points should be similar to line AB. The orientation of line AB is determined by the input weights, while the distance from the origin to line AB is determined by the internal weight originated from the bias neuron. It can be seen that a line passing through the origin (without bias neuron) is not capable of classifying the "+" from the "o".

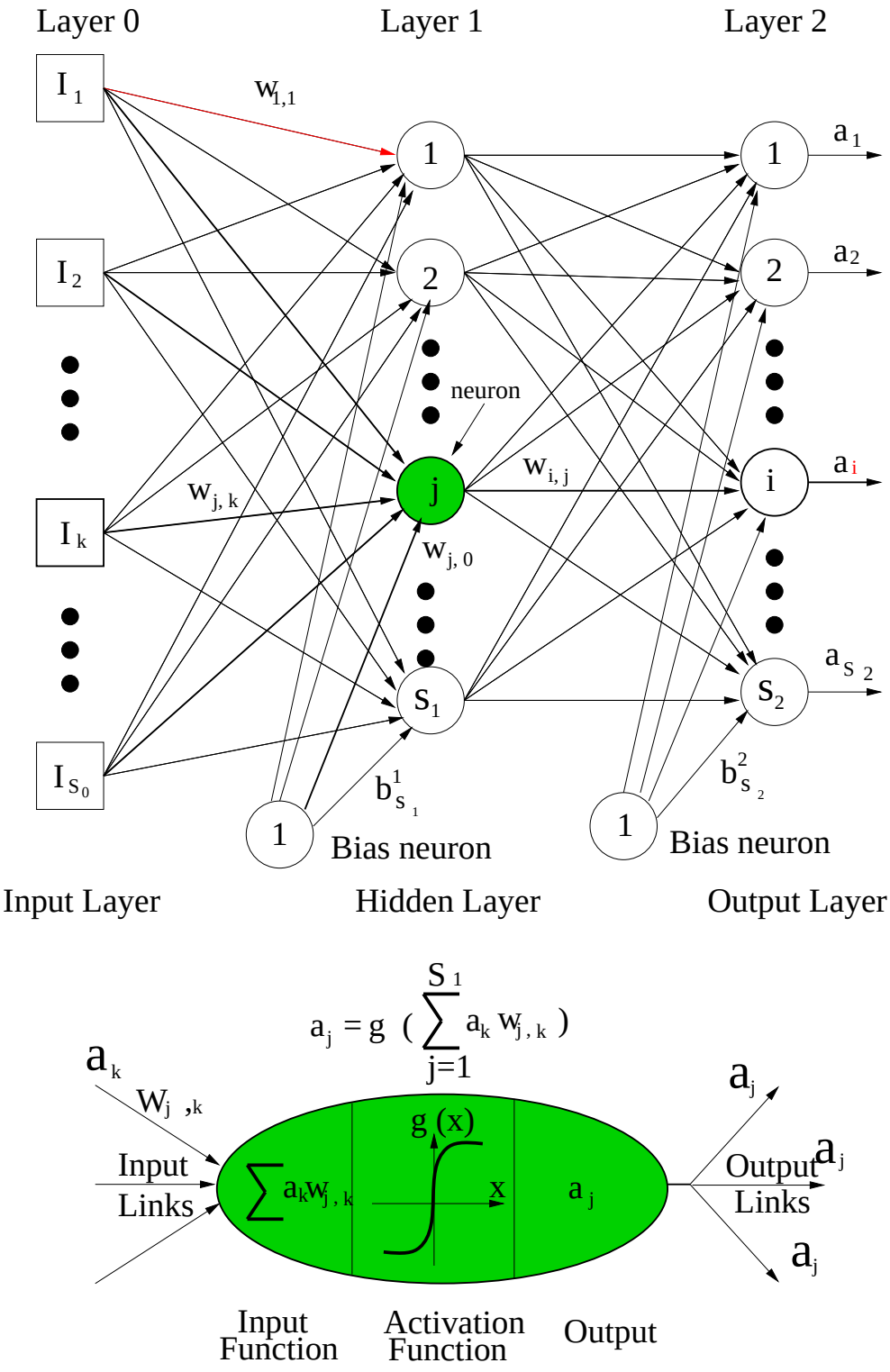

Figure 21: Architecture of Three-layer Feed-Forward Neural Network

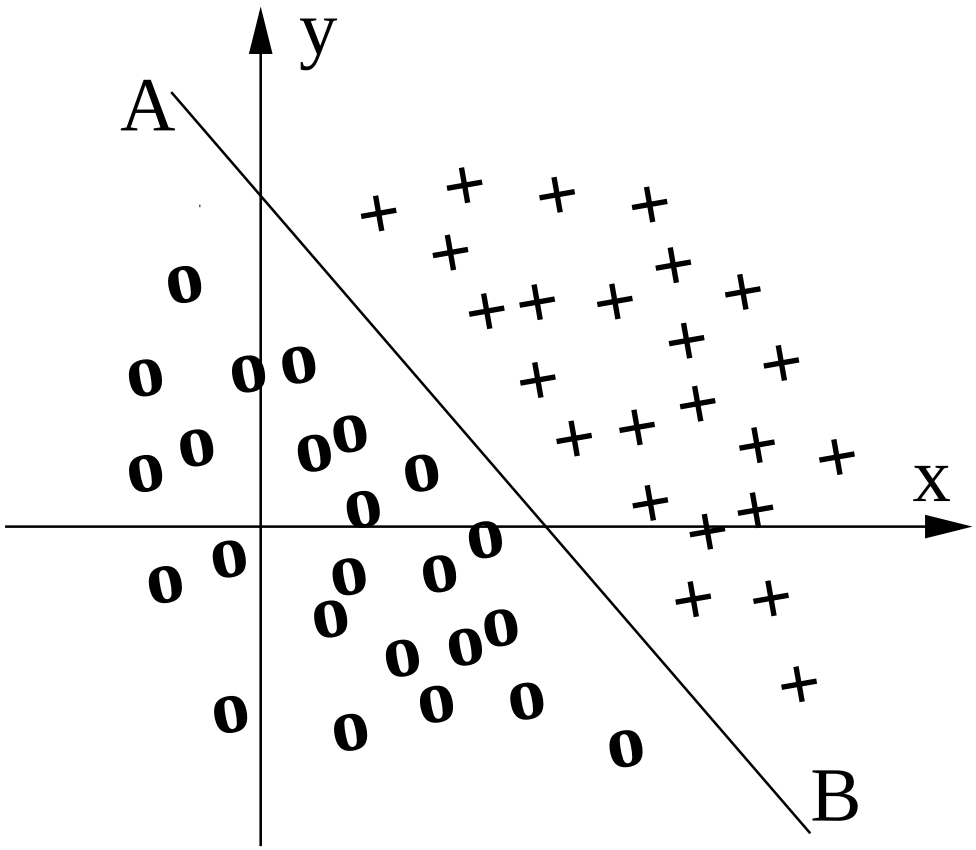

Figure 22: Pattern Classification Example. A neural network is trained to classify the point "+" from the point "o". There is two neurons in the input layer which represent the "x" coordinate and "y" coordinate, one neuron in the output layer and no hidden layer. The output of 1 means point "+" and the output of 0 means point "o". There is a bias neuron in the output layer with fixed output 1. The activation function of the neuron in the output layer is the threshold function which takes the value 1 when the input is greater than zero and 0 otherwise.

5.2 Neural Network Training

With reference to Fig. 21, it can be concluded that the input/output relationships represented by the ANN are determined by its *architecture* (i.e. the number of neurons and the way in which all these neurons are interconnected) as well as the *weights* associated with the neuron connections. Because the architecture of the neural network is typically pre-determined by a "rule of thumb" for a specific problem, the weights associated with the connections are iteratively adjusted so that the ANN produces the desired (i.e. target) output for a given input. This adjustment process is called *training*, and a set of well-chosen input-output pairs which are called the *training data set* is required to complete the training process.

During the training process, the weights are adjusted so that the difference between the output given by the ANN and the target output is minimized for a given training data set. The training process can also be viewed as a minimization process. Like any minimization problem, there is an objective function (or error function) associated with the training process. There are two distinct training modes that correspond to different forms of the objective function. One is called the progressive mode, in which the weights are adjusted based on the ANN's performance on *each* training data pair separately. The other is called the batch mode, in which the weights are adjusted based on the ANN's cumulative performance on *all* training data pairs. The objective function for the batch mode which is used in this project is defined as

$$V = \frac{1}{2}\sum_{q=1}^{Q}\sum_{j=1}^{S_M}(a_{j,q}^{M} - t_{j,q})^2 \tag{44}$$

where Q is the total number of training data pairs, M is the index of the output layer, $t_{j,q}$ is the jth component of the target output vector for qth training data pair, S_M is the number of outputs and $a_{j,q}^{M}$ is the jth component of the neural network's output vector for the qth training data pair.

The algorithm which is commonly used to adjust the weights is the backpropagation (BP) method (Rumelhart, Hinton and Williams, [24]). The essence of the BP method is to adjust each weight in an iterative fashion based on its "contribution" to the objective

function (44). Because the standard BP method typically converges very slowly for large problem, the training methods that are commonly used in practice represent the modified versions of BP such as the BP with momentum (Plaut, Nowlan and Hinton, [25]) and the *quickprop* [26]. In the current project, the BP method with Levenberg-Marquardt algorithm [27] is used.

5.2.1 Backpropagation with Levenberg-Marquardt Method

Upon defining a new vector $\mathbf{x}$ composed of all the interconnection weights and bias in the neural network, the objective function shown in (44) can be rewritten as:

$$V(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{N} e_i^2(\mathbf{x}) \tag{45}$$

where $N = Q \times S_M$, Q is the number of training data pairs, S_M is the number of outputs, and $e_i(\mathbf{x})$ is defined as:

$$e_i(\mathbf{x}) = a_{j,q}^M - t_{j,q}, \qquad i = 1, 2, ..., N \tag{46}$$

where $i = Q(j-1)+q$, $j = 1, 2, ..., S_M$, $q = 1, 2, ..., Q$, $a_{j,q}^M$ is the jth component of the predicted output vector for the qth training data pair, and $t_{j,q}$ is the jth component of the target output vector for the qth training data pair. This form of an objective function is common to general optimization problems and a number of traditional optimization algorithms can be used to minimize it. For example, the gradient search-based iterative algorithm according to traditional Newton's method can be written as

$$\Delta\mathbf{x} = -[\nabla^2 V(\mathbf{x})]^{-1}\nabla V(\mathbf{x}) \tag{47}$$

where $\Delta\mathbf{x}$ is the update of $\mathbf{x}$ resulting from a single iteration, $\nabla^2 V(\mathbf{x})$ is the Hessian operator, and $\nabla V(\mathbf{x})$ is the gradient vector. For the objective function $V(\mathbf{x})$ given by (45), the Hessian matrix and the gradient vector can be computed respectively via

$$\begin{aligned} \nabla^2 V(\mathbf{x}) &= J^T(\mathbf{x})J(\mathbf{x}) + S(\mathbf{x}) \\ \nabla V(\mathbf{x}) &= J^T(\mathbf{x})\mathbf{e}(\mathbf{x}) \end{aligned} \tag{48}$$

where $J(\mathbf{x})$ is the Jacobian matrix, i.e.

$$J(\mathbf{x}) = \nabla \mathbf{e_i}(\mathbf{x}) = \begin{bmatrix} \frac{\partial e_1(\mathbf{x})}{\partial x_1} & \frac{\partial e_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial e_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial e_2(\mathbf{x})}{\partial x_1} & \frac{\partial e_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial e_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e_N(\mathbf{x})}{\partial x_1} & \frac{\partial e_N(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial e_N(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

and $S(\mathbf{x})$ is defined as

$$S(\mathbf{x}) = \sum_{i=1}^{N} e_i(\mathbf{x}) \nabla^2 e_i(\mathbf{x}) \tag{49}$$

To simplify the gradient evaluation following the Gauss-Newton method, it is assumed that $S(\mathbf{x}) \approx 0$ so that

$$\Delta \mathbf{x} \approx -[J^T(\mathbf{x}) J(\mathbf{x})]^{-1} J^T(\mathbf{x}) \mathbf{e}(\mathbf{x}) \tag{50}$$

The Levenberg-Marquardt method is a modification of the standard Gauss-Newton method where:

$$\Delta \mathbf{x} = -[J^T(\mathbf{x}) J(\mathbf{x}) + \mu I]^{-1} J^T(\mathbf{x}) \mathbf{e}(\mathbf{x}) \tag{51}$$

The parameter μ in the Levenberg-Marquardt method plays the role of "exploration" and "stabilization". When an update step increases the objective function, μ is increased by a factor β ($\beta > 1$) to stabilize the minimization process. When a step decreases the objective function, μ is decreased by multiplying $\frac{1}{\beta}$ to explore wider error "surface". In the current project, β takes the value of 10.

Based on the Levenberg-Marquardt method, Hagan and Menhaj ([27]) proposed a fast version of BP method, which is usually called Levenberg-Marquardt method in artificial neural network training. With reference to the symbols defined in Table 9 and Fig. 23, this BP with Levenberg-Marquardt algorithm is summarized as follows:

1. **Calculating Objective Function:**

   All the training pairs are presented to the neural network to evaluate objective function (45) using Eq. (46) based on target output vector and the predicted

Table 9: Symbols used in the Levenberg Marquardt Method

Symbol	Meaning of Symbols
S_k:	Number of neurons in layer k
n_i^k:	Net input of neuron i in layer k
a_i^k:	Output of neuron i in layer k
$\mathbf{a^k}$:	Output vector of layer k, with S_k components
$M+1$:	Total number of layers starting from 0, the input layer is layer 0 and the output layer is layer M
$w_{i,j}^k$:	Weight from neuron j of layer $k-1$ to neuron i in layer k
$\mathbf{W^k}$:	Weight matrix composed of weights from layer $k-1$ to layer k, with dimension $S_{k-1} \times S_k$
b_i^k:	bias of neuron i of layer k
$\mathbf{b^k}$:	bias vector layer k with S_k components
R:	Number of inputs
$\mathbf{I}$:	Input vector with R components
$\mathbf{t}$:	Target output vector with S_M components, given $\mathbf{I}$
f^k:	Activation function for all the neurons in layer k
Q:	Total number of training data pairs

output vector ($\mathbf{a^M}$) obtained through a recursive formula

$$\mathbf{a^0} = \mathbf{I} \tag{52}$$

$$\mathbf{a}^{k+1} = f^{k+1}(\mathbf{W}^{k+1}\mathbf{a}^k + \mathbf{b}^{k+1}) \tag{53}$$

where $k = 0, 1, 2, ..., M-1$, $M+1$ is the total number of layers (including the input layer), f^{k+1} denotes the activation function in layer $k+1$, $\mathbf{W}^{k+1}$ denotes the weighting matrix from layer k to layer $k+1$, $\mathbf{a}^k$ denotes the outputs given by the neurons in layer k, and $\mathbf{b}^{k+1}$ denotes the bias at layer $k+1$. At this time, all weights and biases are kept unchanged.

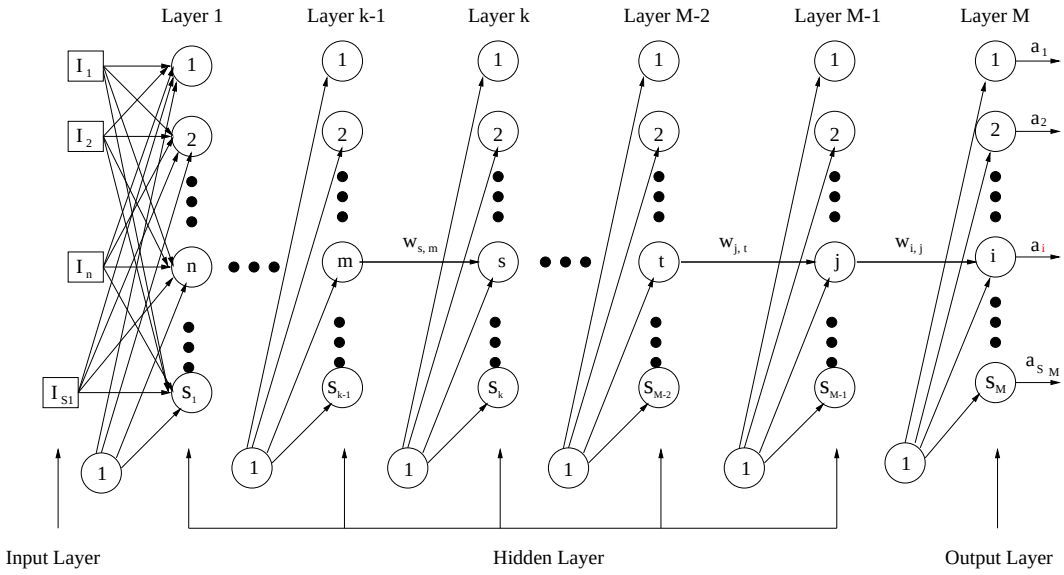

Figure 23: Illustration of Levenberg-Marquardt Algorithm

2. **Calculating Jacobian Matrix:**

Because $\mathbf{x}$ is composed of weights and biases of every layer, different methods should be used for the calculation of the Jacobian matrix. The following equations explain how to calculate the ith row of the Jacobian matrix.

For weight and bias of the output layer ($w_{i,j}^M$ and b_i^M),

$$\begin{aligned}\frac{\partial e_i}{\partial w_{i,j}^M} &= \frac{\partial (t_i - f^M(n_i^M))}{\partial w_{i,j}^M} = -\frac{\partial f^M(n_i^M)}{\partial n_i^M}\frac{\partial n_i^M}{\partial w_{i,j}^M} = -\frac{\partial f^M(n_i^M)}{\partial n_i^M} a_j^{M-1} \\ &\equiv \Delta_i^M a_j^{M-1} \end{aligned} \tag{54}$$

and

$$\begin{aligned}\frac{\partial e_i}{\partial b_i^M} &= \frac{\partial (t_i - f^M(n_i^M))}{\partial b_i^M} = -\frac{\partial f^M(n_i^M)}{\partial n_i^M}\frac{\partial n_i^M}{\partial b_i^M} = -\frac{\partial f^M(n_i^M)}{\partial n_i^M} \\ &= \Delta_i^M \end{aligned} \tag{55}$$

For the weight and bias of layer $M-1$ ($w_{j,t}^{M-1}$ and b_j^{M-1}):

$$\begin{aligned}\frac{\partial e_i}{\partial w_{j,t}^{M-1}} &= \Delta_i^M \frac{\partial f^{M-1}(n_j^{M-1})}{\partial n_j^{M-1}} a_t^{M-2} w_{i,j}^M \\ &\equiv \Delta_t^{M-1} a_t^{M-2} \end{aligned} \tag{56}$$

$$\frac{\partial e_i}{\partial b_j^{M-1}} = \Delta_i^M \frac{\partial f^{M-1}(n_j^{M-1})}{\partial n_j^{M-1}} w_{i,j}^M$$
$$\equiv \Delta_t^{M-1} \tag{57}$$

For the weight and bias of layer k ($0 \leq k \leq M-2$, $w_{s,m}^k$ and b_s^k):

$$\Delta_s^k = \sum_{j=1}^{S_{k+1}} \frac{\partial f^k(n_i^k)}{\partial (n_i^k)} w_{j,i}^{k+1} \Delta_j^{k+1} \tag{58}$$

$$\frac{\partial e_i}{\partial w_{s,m}^k} = \Delta_s^k a_m^{k-1} \tag{59}$$

$$\frac{\partial e_i}{\partial b_s^k} = \Delta_s^k \tag{60}$$

3. **Calculating Weight and Bias Update**

   Solve eq. (51) to obtain $\Delta(\mathbf{x})$.

4. **Recalculate the Objective Function**

   Update the objective function (eq. (45)) using $(\mathbf{x} + \Delta\mathbf{x})$. If $V(\mathbf{x} + \Delta\mathbf{x}) < V(\mathbf{x})$, then multiply μ by $\frac{1}{\beta}$ (decrease μ), update $\mathbf{x}$ by $\mathbf{x} = \mathbf{x} + \Delta\mathbf{x}$ and go back to Step 1 for next iteration. If $V(\mathbf{x} + \Delta\mathbf{x}) \geq V(\mathbf{x})$, then multiply μ by β (increase μ) and go back to Step 3. As stated earlier, $\beta = 10$.

5. **Stop Training**

   The training process is terminated when (i) the value of the objective function falls below the predefined threshold value, or (ii) μ is greater than a predefined large value, or (iii) the predefined maximum number of iterations has been reached, or (iv) the norm of the gradient is less than some predefined value.

5.3 Generalization Ability of ANN

From the statistical point of view, ANN can be seen as a regression model. For any regression tool, an important criterion that evaluates the quality of the model is its generalization ability. An ANN with good generalization ability is capable of producing

good predictions (i.e. output) on novel inputs, which are not included in the training data set. When neural network loses its generality, it gives poor prediction on novel data although it may give very good prediction on the training data.

5.3.1 Over-fitting and Optimal Network Size

One of the reasons that can cause neural network to lose generality is that the network is too large, i.e. that the input-output relationship which the neural network is capable of representing is more complex than necessary. This phenomenon is called *over-fitting*, and it can be demonstrated by training a network to learn the sine function ($y = sin(x)$) in the range $[-2\pi, 2\pi]$. The example network has one neuron in the input layer representing x, one neuron in the output layer representing y, and one hidden layer with n_H neurons. The training data is generated using:

$$y = sin(x) + e \tag{61}$$

where x is an array of equally spaced sampling points within $[-2\pi, 2\pi]$ and e is a Gaussian random variable with mean value equaling zero and variance equaling 1. A total of 41 training data pairs is generated. Fig. 24 shows the comparison of the performance of two networks with $n_H = 4$ and $n_H = 20$. In the diagram, the thick solid line represents the exact sine function, which should be learned by the neural network, while the "+" symbol denotes the "noisy" training data generated using eq. (61). It can be seen that the large network "remembers" the training data set and the noise embedded in it, while the small network "learns" the basic rules embedded in the training data. On the basis of this example, it may be concluded that the ANN should not be too complex for a given input-output relationship.

In solving inverse problems via the ANN approach, however, it is also not appropriate to use excessively small neural network to prevent over-fitting because small neural network is not capable of encoding complex relationships. Furthermore, according to the study of Caruana et.al. [28] and Bartlett [29], the neural network's generalization ability is less sensitive to the size of network than to the magnitude of weights, which

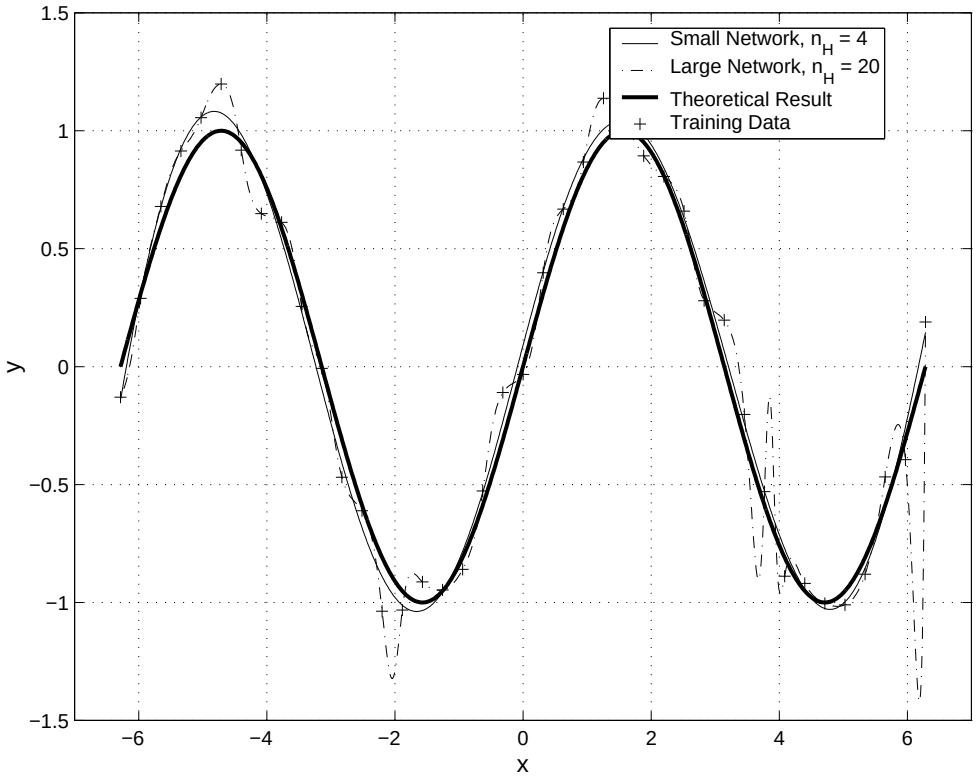

Figure 24: Effect of Network Size on Performance

makes it safe to use a neural network that is slightly more complex than necessary as long as the magnitude of weights do not become too large during training.

For the feed-forward multilayer neural network used in the current project, the adjustable parameters that control the neural network's architecture and therefore its complexity are the (i) number of hidden layers and (ii) number of neurons in each hidden layer. Kim and Kim [8] suggested that the optimal number of neurons (n_H) for a network with one hidden layer should be approximately equal to

$$n_H = round(\frac{n_I + n_O}{2} + \sqrt{n_{train}}) \tag{62}$$

where n_I is the number of inputs, n_O is the number of outputs, n_{train} is the number of training data pairs, and $round(\cdots)$ takes the nearest integer. For the current project where $n_I = 25$ and $n_O = 4$ (as will be explained later), the above recommendation (eq. (62)) implies that the number of weights should approximately be equal to the

number of training data pairs times 0.3. Duda, Hart and Stork ([30]) suggested that the total number of weights should be roughly equal to the number of training data pairs times 0.1. In the current project, networks based on both recommendations are tried, together with an intermediate case in which the number of weights is equal to the number of training data pairs times 0.2. Although the backpropagation network with one hidden layer proved to be capable of approximating any continuous function with sufficient accuracy, a network with two or more hidden layers can achieve similar accuracy with fewer neurons (e.g. [30]). To examine the computational efficiency of different ANNs, neural networks having one and two hidden layers with the above number of weights were both trained.

5.3.2 Overtraining and Early Stopping

In neural network training, the magnitude of individual weights can become excessively large when the neural network is *overtrained*. In general, overtraining results in similar phenomena (i.e. loss of the generalization ability) as over-fitting. However, while over-fitting relates to the inherent excessive complexity of the neural network, overtraining is caused by over-minimization of the objective function which forces the neural network to *remember* the training data rather than *learn* the rules embedded in the training data.

On the other hand, the neural network will not learn the necessary rules well if the objective function is not minimized to an acceptable level. According to Caruana et.al. ([28]), a technique called *early stopping* can overcome this problem. In this method, a *separate* set of input-output data pairs is used to monitor the performance of the trained network. This separate set of input-output data pairs, which must be novel to the neural network (i.e. distinct from the training data set), is usually called *test data set*. When the prediction error on the test data set reaches its minimum value and begins to increase, the training is terminated and the network corresponding to the minimum prediction error on test data is taken as the optimal neural network (Duda, Hart and Stork, [30]). This is another criterion on stopping training described in Section 5.2.1.

5.3.3 Noise Injection

Because the neural network developed in this project is designed to be used on the field data in order to predict the pavement properties, it is logical to expect that the experimental input to ANN would contain some variability. This is shown in Fig. 9, which plots the frequency response function from multiple FWD tests performed on the same test location on the same day. From the Figure, one can see that the FWD measurements have significant variance even for the same test location. To develop a trained neural network that is robust in the presence of such noisy inputs (i.e., to ensure that the ANN has a good generalization ability on noisy inputs), one approach is to add some variability (i.e. perturbation) into the input portion of the training data set, while keeping the corresponding output unchanged (Matsuoka, [31], Grandvalet and Canu [32]). This technique is called *noise injection*. In this study, the training data are generated by simulating the FWD test using a three dimensional multilayer viscoelastic dynamic model that produces results with no variability, i.e. noise-free training data. Upon denoting the noise-free training data pair as ($\mathbf{I}'$, $\mathbf{t}'$) where $\mathbf{I}'$ is the input vector and $\mathbf{t}'$ is the output vector, the "noisy" training data pair can be written as ($\mathbf{I}$, $\mathbf{t}$) with its components defined using either

$$\begin{aligned} I_k &= I'_k(1 + \alpha_{rel} R_k) \\ t_j &= t'_j \end{aligned} \tag{63}$$

or

$$\begin{aligned} I_k &= I'_k + \alpha_{abs} R_k \\ t_j &= t'_j \end{aligned} \tag{64}$$

where I_k and t_j are the components of the input vector and output vector of "noisy" training data pair, while I'_k and t'_j are the components of the input and output vector comprising the noise-free training data pair. In the above, R is a random variable with uniform distribution within $[-1, 1]$, α_{rel} is the relative noise level, and α_{abs} denotes the

absolute noise level. For a given noise-free training data pair ($\mathbf{I}'$, $\mathbf{t}'$), eq. (63) and/or eq. (64) is applied to all components of the exact (i.e. noise-free) input vector N_{rep} times. For example, if there are 5000 noise-free training data pairs, there will be $N_{rep} \times 5000$ noisy training data pairs used in the neural network training. In the following, the expression "number of training data pairs" will refer to the number of noise-free training data pairs.

6 Backcalculation Using ANN

6.1 Back-analysis Approach

Based on the foregoing developments (see Fig. 4), the ANN based backcalculation scheme developed in this project is detailed in Fig. 25. This scheme is explained as follows:

- For a given type of hot-mix asphalt at certain temperature, its viscosity is calculated using eqs. (16). Together with the HMA grading parameters, this viscosity is applied to the empirical equation (15) (Witczak [17]) to calculate the complex Young's modulus of the HMA. Second, the parameters in the power law model (D_0, D_1 and m) in eq. (7) are adjusted using regression so that the amplitude of the complex Young's modulus given by the power law model (10) is approximately equal to that given by empirical equation (15). (Fig. 26 shows the example comparison of the amplitude of the complex Young's modulus given by empirical formula and that given by the power law model.) Finally, the complex shear modulus (G^*_{HMA}) and the complex Poisson's ratio (ν^*_{HMA}) of the asphalt layer are calculated using eqs. (10) to (14).

- Thickness of the HMA layer (h_{HMA}) and aggregate base (h_B) are obtained from the construction record.

- Frequency response functions (FRF) of interest are estimated from the FWD measurements at geophones 2, 5 and 8 using eq. (30).

- Parameters of the power law model, the thickness of HMA, the thickness of the aggregate base, and the segmental areas under the real and imaginary components of the selected frequency response functions comprise an input to the ANN (see Fig. 25). The outputs of the ANN are the thickness of subgrade (h_S) and the shear moduli of aggregate base, subgrade and half-space (G_B, G_S and G_H).

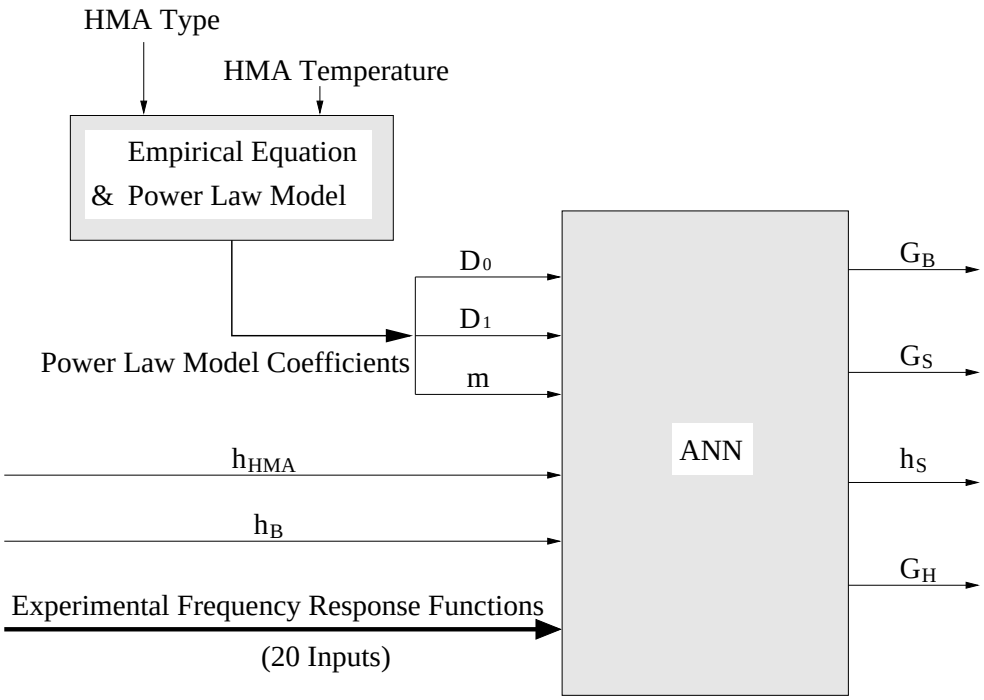

Figure 25: ANN Based Backcalculation Model

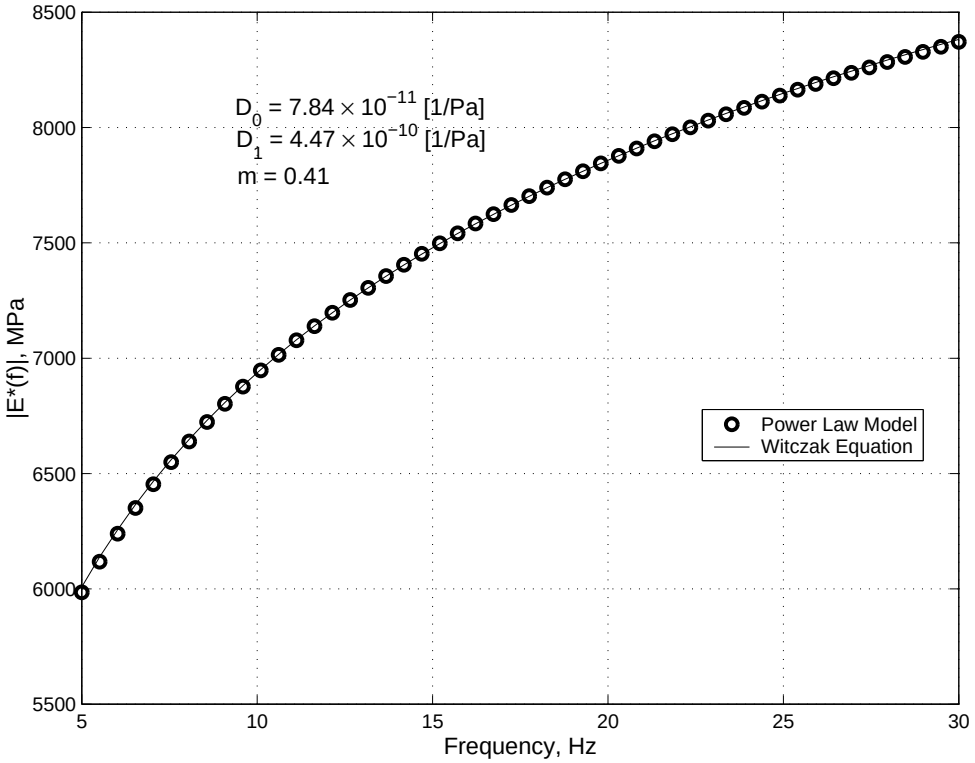

Figure 26: Comparison of $|E^*|$ given by Empirical Formula (Witczak's equation) and $|E^*|$ given by Fitted Power Law Model for PG 58-28 at 9.1^o (Celsius)

6.2 Generation of the Training and Test Data Sets

As stated before, the input-output relationship embedded in the neural network is obtained by training the neural network using a properly selected, synthetic training data set. In the current project, the training data set is generated by simulating the FWD test using a computer program, which is an implementation of the three-dimensional multilayer viscoelastic dynamic model (Guzina and Pak, [4]). With reference to Tables 6 and 7, the range of pavement properties that are used to generate the training data set are summarized in Tables 10 and 11 that list (i) the parameters whose values are assumed to be fixed, and (ii) the ones with variate values, respectively.

Table 10: Pavement Properties with Fixed Values

Property	h [m]	G [MPa]	ν	ρ [kg/m^3]	ξ_s	ξ_p
HMA	Variable Value	N/A	0.4	2335	N/A	N/A
Aggregate Base	Variable Value	Variable Value	0.35	2027	0.01	0.005
Subgrade	Variable Value	Variable Value	0.4	1865	0.01	0.005
Half-space	N/A	Variable Value	0.45	2160	0.01	0.005

Table 11: Pavement Properties with Variate Values

Property	h [m]	G [MPa]	E [MPa]
HMA	$0.1 \sim 0.4$	N/A	$1000 \sim 25000$
Aggregate Base	$0.1 \sim 1.0$	$10 \sim 130$	$27 \sim 351$
Subgrade	$0.3 \sim 3.0$	$10 \sim 130$	$28 \sim 364$
Half-space	N/A	$10 \sim 130$	$29 \sim 377$
Property	D_0 [10^{-11}Pa^{-1}]	D_1 [10^{-11}Pa^{-1}]	m
HMA	$4 \sim 100$	$2 \sim 700$	$0.15 \sim 0.45$

On the basis of Tables 10 and 11, the training data set is generated using the follow-

ing procedure, where R is a random variable with uniform distribution within $[-1, 1]$, $[f_{min}, f_{max}]$ is the selected frequency range ($f_{min} = 1$ Hz and $f_{min} = 51$ Hz based on eq. 26), $N_f = 50$ is the total number of sampling frequencies, and $f_{cal,j}$ $(j = 1, 2, ..., N_f)$ is the sampling frequency defined as

$$f_{cal,j} = f_{min} + (f_{max} - f_{min})\frac{j-1}{N_f - 1} \tag{65}$$

1. Set the values of the fixed pavement parameters according to Table 10:

$$\begin{aligned} \rho_{HMA} &= 2335 \text{ kg/m}^3 \\ \rho_B &= 2027 \text{ kg/m}^3 \\ \rho_S &= 1865 \text{ kg/m}^3 \\ \rho_H &= 2160 \text{ kg/m}^3 \\ \nu_{HMA} &= 0.4 \\ \nu_B &= 0.35 \\ \nu_S &= 0.4 \\ \nu_H &= 0.45 \\ \xi_{s,B} &= 0.01 \\ \xi_{s,S} &= 0.01 \\ \xi_{s,H} &= 0.01 \\ \xi_{p,B} &= 0.005 \\ \xi_{p,S} &= 0.005 \\ \xi_{p,H} &= 0.005 \end{aligned} \tag{66}$$

2. Initialize the index of the generated input/output training data pair (k) to zero;

3. With reference to Table 11, specify the *randomly-selected* values of the variate pavement parameters according to:

$$h_{HMA} = 0.1 + (0.4 - 0.1)(0.5R + 0.5) \quad \text{[m]} \tag{67}$$

$$h_B = 0.1 + (1.0 - 0.1)(0.5R + 0.5) \quad [\text{m}]$$

$$h_S = 0.3 + (3.0 - 0.3)(0.5R + 0.5) \quad [\text{m}]$$

$$G_B = 10 + (130 - 10)(0.5R + 0.5) \quad [\text{MPa}]$$

$$G_S = 10 + (130 - 10)(0.5R + 0.5) \quad [\text{MPa}]$$

$$G_H = 10 + (130 - 10)(0.5R + 0.5) \quad [\text{MPa}]$$

$$D_0 = 4 \times 10^{-11} + (100 \times 10^{-11} - 4 \times 10^{-11})(0.5R + 0.5) \quad [1/\text{Pa}]$$

$$D_1 = 2 \times 10^{-11} + (700 \times 10^{-11} - 2 \times 10^{-11})(0.5R + 0.5) \quad [1/\text{Pa}]$$

$$m = 0.15 + (0.45 - 0.15)(0.5R + 0.5)$$

while satisfying the constraints:

$$G_B > G_S \qquad \text{and} \qquad G_S < G_H \tag{68}$$

The above constraint is enforced because the aggregate base is typically stiffer than the subgrade soil, and the half-space (stiff layer) is usually saturated or frozen which makes it stiffer than the subgrade.

4. Set the index for calculation frequency j to 1

5. For frequency $f_{cal,j}$, calculate the frequency-dependent parameters $G^*_{HMA}(f_{cal,j})$ and $\nu^*_{HMA}(f_{cal,j})$ according to eqs. (12) and (13)

6. Substitute $G^*_{HMA}(f_{cal,j})$, $\nu^*_{HMA}(f_{cal,j})$ and parameters set by eqs. (66) and (67) into the forward model to calculate the displacement response for geophones 2, 5 and 8 at frequency $f_{cal,j}$, namely, $D_2(f_{cal,j})$, $D_5(f_{cal,j})$ and $D_8(f_{cal,j})$.

7. Increase the index for calculation frequency j by one, if j is greater than N_f, continue to Step 8, otherwise go to Step 5.

8. Use the calculated displacement response at geophones 2, 5 and 8 within $[f_{min}, f_{max}]$ to calculate the frequency response functions according to eq. (22) which are

rewritten here for convenience:

$$FRF_{25} = \frac{D_5(f)}{D_2(f)} \qquad (69)$$
$$FRF_{58} = \frac{D_8(f)}{D_5(f)}$$

9. Use the integration scheme stated in Section 5.1.5, integrate the real and imaginary parts of FRF_{25} and FRF_{58} over the five selected frequency intervals, namely $[f_1, f_2]$, $[f_2, f_3]$, $[f_3, f_4]$, $[f_4, f_5]$ and $[f_5, f_6]$, where:

$$f_1 = 1\text{Hz}$$
$$f_2 = 11\text{Hz}$$
$$f_3 = 21\text{Hz}$$
$$f_4 = 31\text{Hz}$$
$$f_5 = 41\text{Hz}$$
$$f_6 = 51\text{Hz}$$

With reference to Fig. 12b, the resulting segmental integrals are denoted by $Area^{kj}_{real,25}$, $Area^{kj}_{imag,25}$, $Area^{kj}_{real,58}$ and $Area^{kj}_{imag,58}$, where $k = 1, 2, 3, 4, 5$, and $j = k + 1$ conrrespondingly.

10. Based on the method of noise injection described in Section 6.3.3, construct N_{rep} noisy input/output pairs by applying eq. (63) and/or eq. (64) N_{rep} times to the noise-free segmental areas obtained in Step 9. The resulting (i.e. noise-contaminated) input vector ($\mathbf{I}$) and output vector ($\mathbf{t}$) can be summarized as

$$\mathbf{I} = \begin{bmatrix} (1+\alpha_{rel}R)Area^{12}_{real,25} \\ (1+\alpha_{rel}R)Area^{12}_{imag,25} \\ (1+\alpha_{rel}R)Area^{23}_{real,25} \\ (1+\alpha_{rel}R)Area^{23}_{imag,25} \\ (1+\alpha_{rel}R)Area^{34}_{real,25} \\ (1+\alpha_{rel}R)Area^{34}_{real,25} \\ (1+\alpha_{rel}R)Area^{45}_{imag,25} \\ (1+\alpha_{rel}R)Area^{45}_{real,25} \\ (1+\alpha_{rel}R)Area^{56}_{imag,25} \\ (1+\alpha_{rel}R)Area^{56}_{real,25} \\ (1+\alpha_{rel}R)Area^{12}_{real,58} \\ (1+\alpha_{rel}R)Area^{12}_{imag,58} \\ (1+\alpha_{rel}R)Area^{23}_{real,58} \\ (1+\alpha_{rel}R)Area^{23}_{imag,58} \\ (1+\alpha_{rel}R)Area^{34}_{real,58} \\ (1+\alpha_{rel}R)Area^{34}_{real,58} \\ (1+\alpha_{rel}R)Area^{45}_{imag,58} \\ (1+\alpha_{rel}R)Area^{45}_{real,58} \\ (1+\alpha_{rel}R)Area^{56}_{imag,58} \\ (1+\alpha_{rel}R)Area^{56}_{real,58} \\ (1+\alpha_{rel}R)D_1 \text{ or } D_1+\alpha_{abs}R \\ (1+\alpha_{rel}R)D_0 \text{ or } D_0+\alpha_{abs}R \\ (1+\alpha_{rel}R)m \text{ or } m+\alpha_{abs}R \\ (1+\alpha_{rel}R)h_{HMA} \text{ or } h_{HMA}+\alpha_{abs}R \\ (1+\alpha_{rel}R)h_B \text{ or } h_B+\alpha_{abs}R \end{bmatrix} \tag{70}$$

$$\mathbf{t} = \begin{bmatrix} G_B \\ G_S \\ h_S \\ G_H \end{bmatrix} \tag{71}$$

where α_{rel} denotes the relative noise level, and α_{abs} denotes the absolute noise level. The selection of α_{rel} and α_{abs} for the last five components in $\mathbf{I}$ will be discussed in detail in Section 6.3.

11. Increase the index of the generated input/output training data pair (k) by one. If k is greater than the total number of training data pairs required, the computation terminates; otherwise, go to Step 3.

6.3 ANN Development

From the foregoing outline of the backcalculation technique proposed (see Fig. 25), it can be seen that the neural network is the key component of the featured method. To develop an optimal ANN for the problem of interest, many alternative neural networks with different architectures were developed. As stated before, the ANN-based backcalculation is designed to be applied to actual FWD measurements as a tool for estimating the stiffness properties of the pavement section. To aid the development process, backcalculation models based on different ANNs are applied to actual FWD measurements obtained at the Minnesota Road Research Project (Mn/ROAD) testing facility, and the results are studied to decide which ANN model is most appropriate for the inverse problem of interest. In this process, competing neural networks are trained using the Matlab Neural Network Toolbox ([22]) and numerous Matlab routines written by the author.

6.3.1 Summary of Neural Networks Developed

To find the optimal neural network, a total of 32 ANNs were developed based on the considerations given in Section 5.3. For brevity, only five representative neural networks will be discussed. They are described as follows.

- **ANN 1** There are (i) 25 inputs to the network including 20 FRF values, D_0, D_1, m, h_{HMA} and h_B, and (ii) 4 outputs including G_B, G_S, h_S and G_H (see Fig. 27). The network has two hidden layers, with 40 neurons in the first hidden layer and 20 neurons in the second hidden layer. The network is trained using $10,000$ *training* data pairs and 653 *test* data pairs. In the noise injection process (see eq. (70)), the relative noise level α_{rel} which represents the perturbation of the noise-free FRF's is taken as $\alpha_{rel} = 0.03$, while the noise level for the remaining five input components takes the form of relative noise level with a value of $\alpha_{rel} = 0.03$ (eq. (63)).

- **ANN 2** There are (i) 25 inputs to the network including 20 FRF values, D_0, D_1, m, h_{HMA} and h_B, and (ii) 4 outputs including G_B, G_S, h_S and G_H (see Fig. 27).

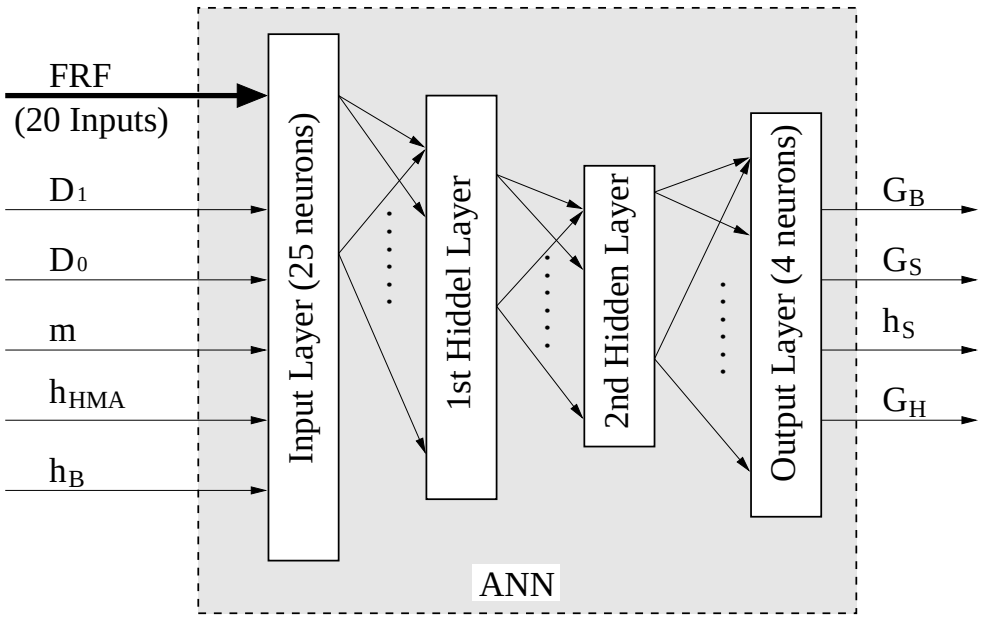

Figure 27: Illustration of ANN 1, 2, 3 and 4

The network has two hidden layers, with 24 neurons in the first hidden layer and 12 neurons in the second hidden layer. The network is trained using $5,000$ *training* data pairs and 653 *test* data pairs. In the noise injection process (see eq. (70)), the relative noise level α_{rel} which represents the perturbation of the noise-free FRF's is taken as $\alpha_{rel} = 0.2$, while the noise level for the remaining five input components takes the form of relative noise level with a value of $\alpha_{rel} = 0.2$ (eq. (63)).

- **ANN 3** There are (i) 25 inputs to the network including 20 FRF values, D_0, D_1, m, h_{HMA} and h_B, and (ii) 4 outputs including G_B, G_S, h_S and G_H (see Fig. 27). The network has two hidden layers, with 40 neurons in the first hidden layer and 20 neurons in the second hidden layer. The network is trained using $10,000$ *training* data pairs and 653 *test* data pairs. In the noise injection process (see eq. (70)), the relative noise level α_{rel} which represents the perturbation of the noise-free FRF's is taken as $\alpha_{rel} = 0.2$; the noise level for D_0, D_1 and m takes the form of relative noise level with a value of $\alpha_{rel} = 0.03$ (eq. (63)); the noise level for h_{HMA} takes the absolute value $\alpha_{abs} = 0.0127$ meters, and the noise level for h_B takes the absolute

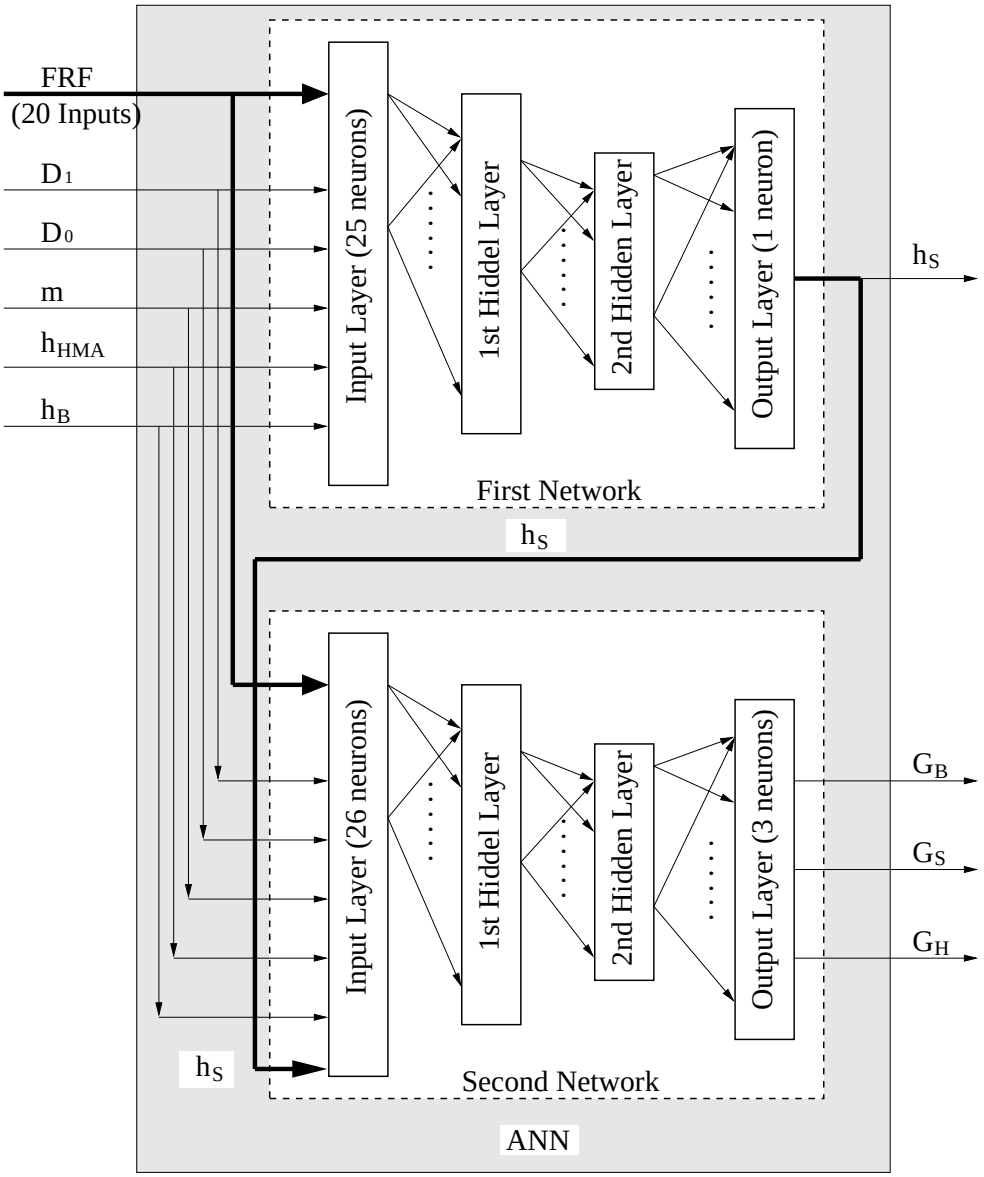

Figure 28: Illustration of ANN 5

value $\alpha_{abs} = 0.0254$ meters (eq. (64)).

- **ANN 4** There are (i) 25 inputs to the network including 20 FRF values, D_0, D_1, m, h_{HMA} and h_B, and (ii) 4 outputs including G_B, G_S, h_S and G_H (see Fig. 27). The network has two hidden layers, with 24 neurons in the first hidden layer and 12 neurons in the second hidden layer. The network is trained using $10,000$ *training*

data pairs and 653 *test* data pairs. In the noise injection process (see eq. (70)), the relative noise level α_{rel} which represents the perturbation of the noise-free FRF's is taken as $\alpha_{rel} = 0.2$; the noise level for D_0, D_1 and m takes the form of relative noise level with a value of $\alpha_{rel} = 0.03$ (eq. (63)); the noise level for h_{HMA} takes the absolute value $\alpha_{abs} = 0.0127$ meters, and the noise level for h_B takes the absolute value $\alpha_{abs} = 0.0254$ meters (eq. (64)).

- **ANN 5 (ANN Committee)** In what follows, this model will also be referred to as the sequential approach since the two consecutive neural networks are used to solve the problem.

  In the first network, there are (i) 25 inputs to the network including 20 FRF values, D_0, D_1, m, h_{HMA} and h_B, and (ii) 1 output, namely h_S (see Fig. 28). The network has two hidden layers, with 40 neurons in the first hidden layer and 20 neurons in the second hidden layer. The network is trained using $10,000$ *training* data pairs and 653 *test* data pairs. In the noise injection process (see eq. (70)), the relative noise level α_{rel} which represents the perturbation of the noise-free FRF's is taken as $\alpha_{rel} = 0.2$; the noise level for D_0, D_1 and m takes the form of relative noise level with a value of $\alpha_{rel} = 0.03$ (eq. (63)); the noise level for h_{HMA} takes the absolute value $\alpha_{abs} = 0.0127$ meters, and the noise level for h_B takes the absolute value $\alpha_{abs} = 0.0254$ meters (eq. (64)).

  In the second network, there are (i) 26 inputs to the network including 20 FRF values, D_0, D_1, m, h_{HMA}, h_B and h_S which is the output of the first network, and (ii) 3 outputs including G_B, G_S and G_H (see Fig. 28). The network has two hidden layers, with 40 neurons in the first hidden layer and 20 neurons in the second hidden layer. The network is trained using $10,000$ *training* data pairs and 653 *test* data pairs. In the noise injection process (see eq. (70)), the relative noise level α_{rel} which represents the perturbation of the noise-free FRF's is taken as $\alpha_{rel} = 0.2$; the noise level for D_0, D_1 and m takes the form of relative noise level with a value of $\alpha_{rel} = 0.03$ (eq. (63)); the noise level for h_{HMA} takes the

absolute value $\alpha_{abs} = 0.0127$ meters, the noise level for h_B takes the absolute value $\alpha_{abs} = 0.0254$ meters, and the noise level for h_S takes the absolute value $\alpha_{abs} = 0.3$ meters, which is the averaged prediction error of the first network (eq. (64)).

6.3.2 Computation Requirements

The neural network development is in general a computationally extensive process. In this project, development of the ANN took approximately six months. The reasons for such high computational cost are elucidated in the following.

First, training data generation took a significant computation time because the FWD is simulated using a three dimensional dynamic model. To generate one training data pair, it takes approximately 4 minutes. Including the $10,000$ training data pairs which were used to train the final network, a total of $28,000$ training data have been generated, which would take about 78 days on a single personal computer. To accelerate this process, two to three computers were used simultaneously over the period of one month.

Second, because of the large size of neural network used in this project, the neural network training requires a long period of time. To accelerate the training, the second-order Levenberg-Marquardt algorithm is used. However, the price of this acceleration is the requirement for a large amount of physical memory, which is 7 to 12 GB for the neural network developed. For each neural network, it takes an average of two days to complete the training process.

Third, to choose a neural network that is most appropriate for the FWD field measurement, all neural networks developed were applied to field FWD measurements. It takes about 4 minutes to apply a neural network to one measurement (i.e., specific combination of specific test time, test section, load level and test location). For all available FWD field measurements, this process takes about 400 hours.

7 Results

7.1 Performance on Synthetically-generated Test Data

To assess their performance, the competing backcalculation models based on the above five ANN's were applied to the synthetically-generated test data set first. In Figs. 29 to 33, the predicted values of each output parameter (G_B, G_S, h_S and G_H) as given by the ANN backcalculation model are plotted against their true values given by the test data set. In the ideal case, the predicted values should be the same as their true values for all test data, and all the points in the figure would lie on the line with the 45 degree slope denoted by the thick line in the figures. Also shown in the figure is the correlation coefficient defined as:

$$r = \frac{Cov(\mathbf{P}, \mathbf{T})}{\sqrt{Cov(\mathbf{P}, \mathbf{P})}\sqrt{Cov(\mathbf{T}, \mathbf{T})}} \tag{72}$$

where $\mathbf{P}$ represents the vector of predicted values of the selected property, $\mathbf{T}$ represents the vector of true values of the same property, and $Cov(\mathbf{P}, \mathbf{T})$ is the covariance between $\mathbf{P}$ and $\mathbf{T}$. In the ideal case, the correlation coefficient is equal to unity.

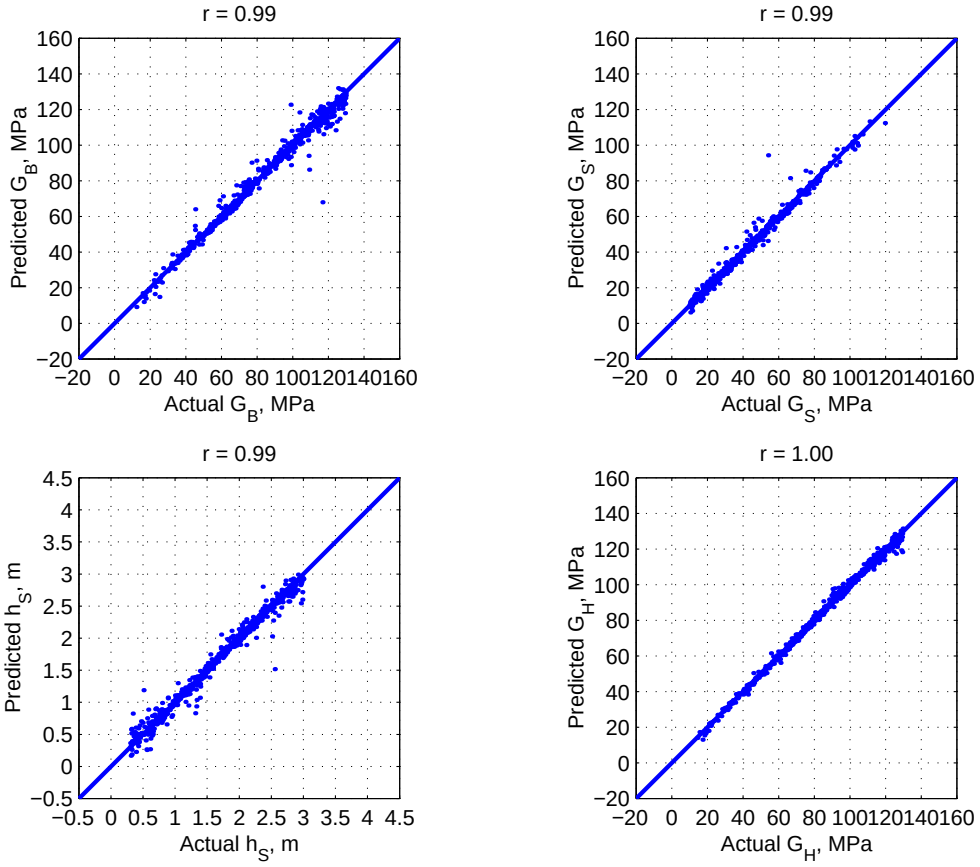

Figure 29: ANN Performance on Synthetically-generated Test Data, ANN 1

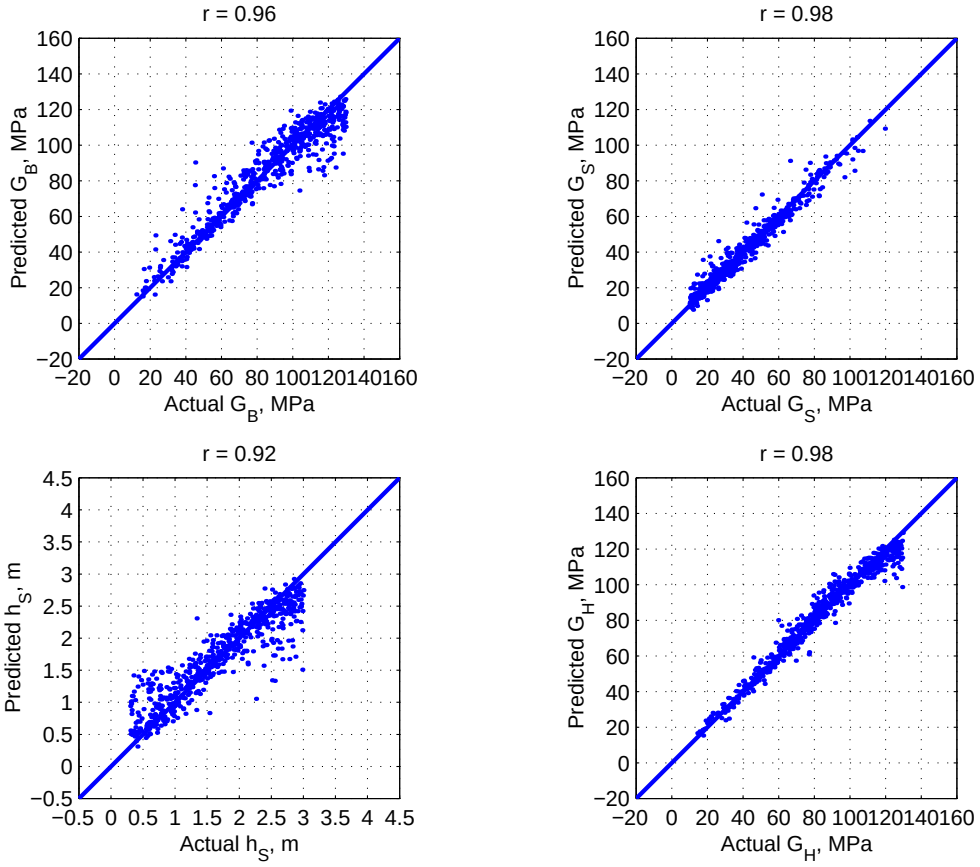

Figure 30: ANN Performance on Synthetically-generated Test Data, ANN 2

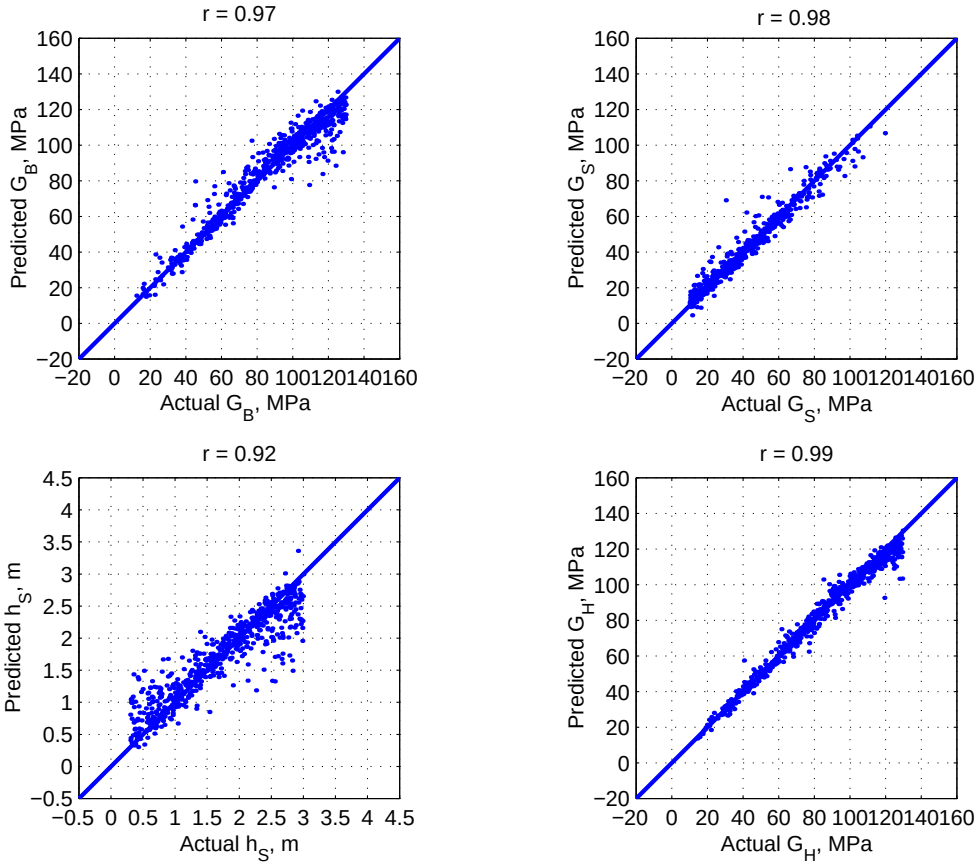

Figure 31: ANN Performance on Synthetically-generated Test Data, ANN 3

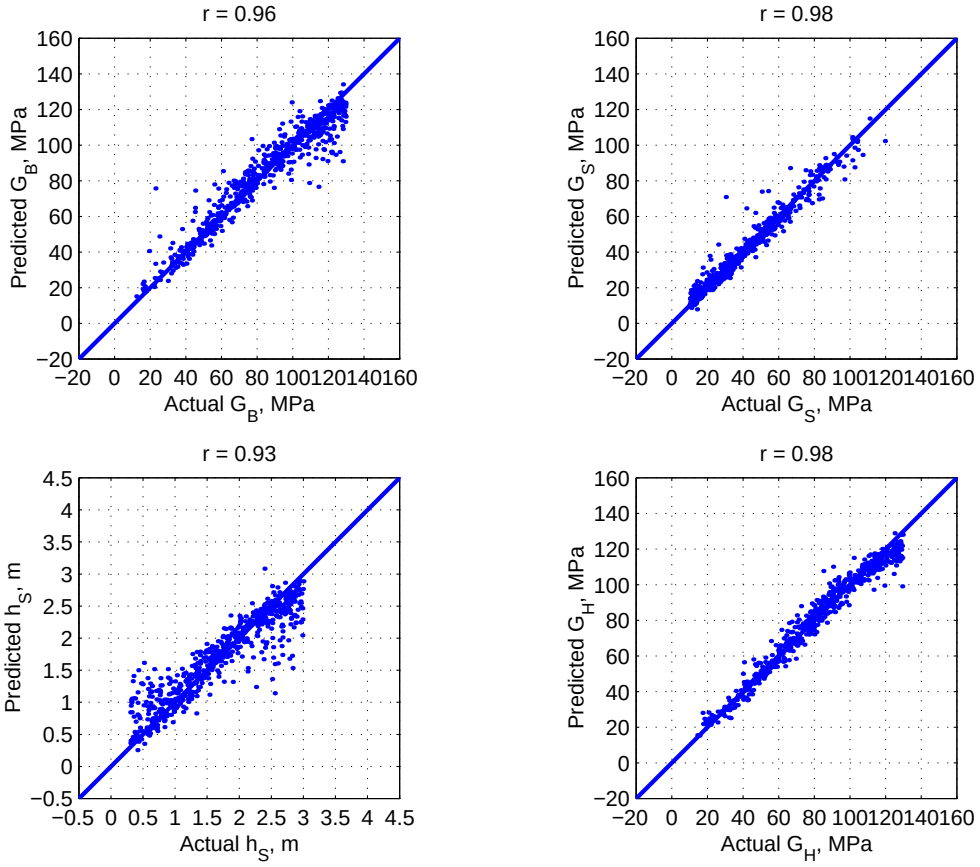

Figure 32: ANN Performance on Synthetically-generated Test Data, ANN 4

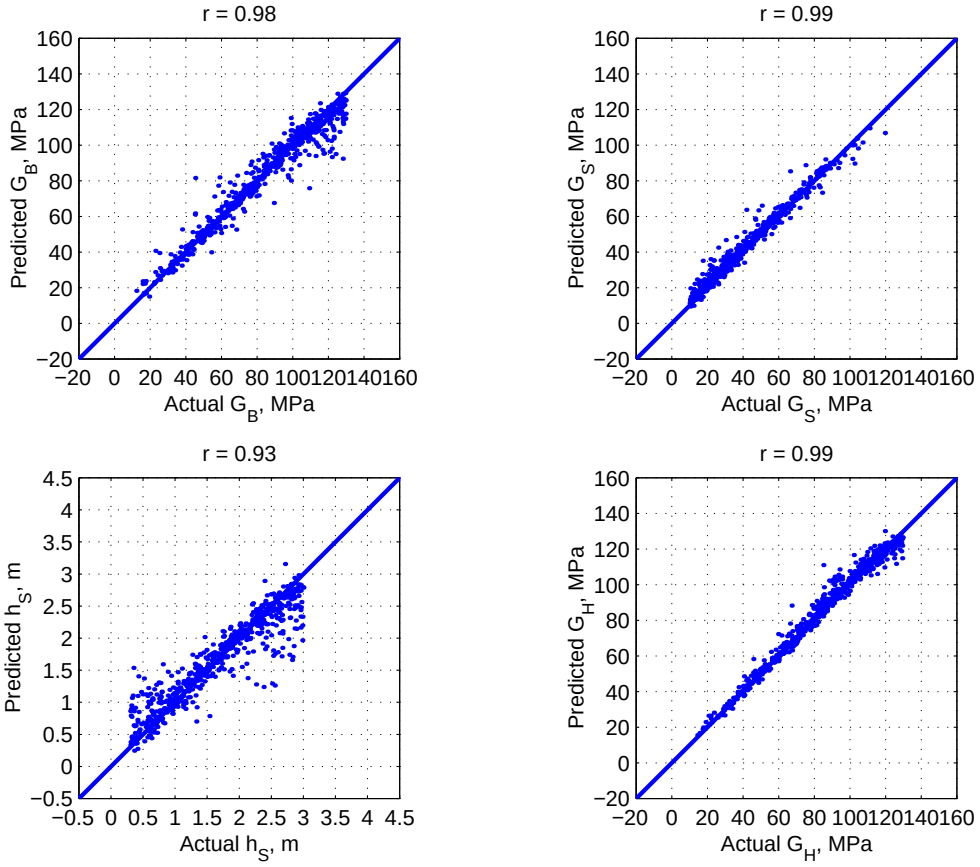

Figure 33: ANN Performance on Synthetically-generated Test Data, ANN 5

7.1.1 Prediction Error

On the basis of the results plotted in Figs. 29 to 33, Table 12 summarizes the averaged prediction error (L_2 norm) on each property for all five neural networks. From the

Table 12: ANN Prediction Error (L_2 norm)

ANN	G_B [MPa]	G_S [MPa]	h_S [m]	G_H [MPa]
ANN 1	4	3	0.10	2
ANN 2	9	4	0.31	5
ANN 3	7	4	0.31	4
ANN 4	8	4	0.30	5
ANN 5	7	4	0.28	4

table, it can be seen that the average prediction error for the thickness of subgrade is approximately 0.3 meters, except for ANN 1. Such prediction error can be justified by the inherent resolution limits of seismic waves. According to the seismological literature (Farr [10]), seismic waves are typically not capable of resolving layers thinner than $\lambda_{min}/12$, where λ_{min} is the minimum wave length. To illustrate, the minimum shear wave length in the subgrade layer for the problem of interest can be calculated as:

$$\lambda_{min} = \frac{c_{s,S}}{f_{max}} \qquad (73)$$
$$c_{s,S} = \sqrt{\frac{G_S}{\rho_S}}$$

where $c_{s,S}$ is the shear wave speed of the subgrade, G_S and ρ_S are the respective shear modulus and mass density of the subgrade, and $f_{max} = 51$ Hz is the upper frequency limit of FWD measurements used in the current project. Substituting $G_S = 70$ MPa (which is an average subgrade shear modulus from Table 11) and $\rho_S = 1865$ kg/m^3 into eq. (73), the subgrade thickness resolution can be calculated as

$$\frac{\lambda_{min}}{12} = 0.31m \qquad (74)$$

which compares well with the average prediction error on h_S given in Table 12.

From Table 12 and Figs. 29 to 33, it can also be seen that relative to Artificial Neural Networks 2, 3, 4 and 5, ANN 1 has a lower level of noise in the training data set and thus has a better performance on the test data set. As a result, from the stand point of noise-free test data, higher noise level leads to the drop in ANN performance.

7.2 Performance on Actual FWD Measurements

As stated earlier, the ANN backcalculation model developed in this study is designed to be applied to field FWD measurements. As a result, the prediction performance on the actual FWD measurements is considered to be the most appropriate criterion for evaluating the effectiveness of each neural network. There are two representative low-volume road (LVR) test sections at the Mn/ROAD research facility (see Fig. 34) from which the FWD test data were used to evaluate the foregoing five neural networks.

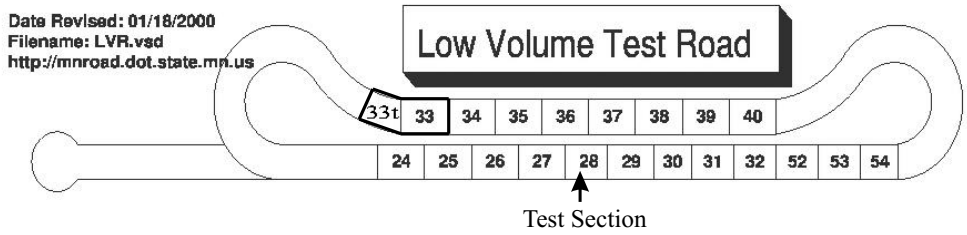

Figure 34: Low Volume Road (LVR) Layout (Excerpted from http://www.mrr.dot.state.mn.us /research/MnROAD_Project/LVR.pdf and Modified by the author)

7.2.1 Regular Pavement Section

The backcalculation methodology based on the foregoing neural networks is applied to the representative FWD measurements performed at the Mn/ROAD Test Section 33 (see Fig. 34) on May 22, 2001, for which the testing locations are shown in Fig. 35.

For this test, Table 13 shows the predicted pavement properties as given by each of the five neural networks, and for each ANN, the properties shown in the Table are the average values over the prediction results on six testing locations. Because the FWD measurements in the current project are interpreted in terms of the frequency response functions, it is also natural to compare the prediction quality of above five ANNs by comparing the frequency response function (FRF's) calculated using the estimated pavement properties (G_B, G_S, h_S and G_H) with the experimental FRF's computed from

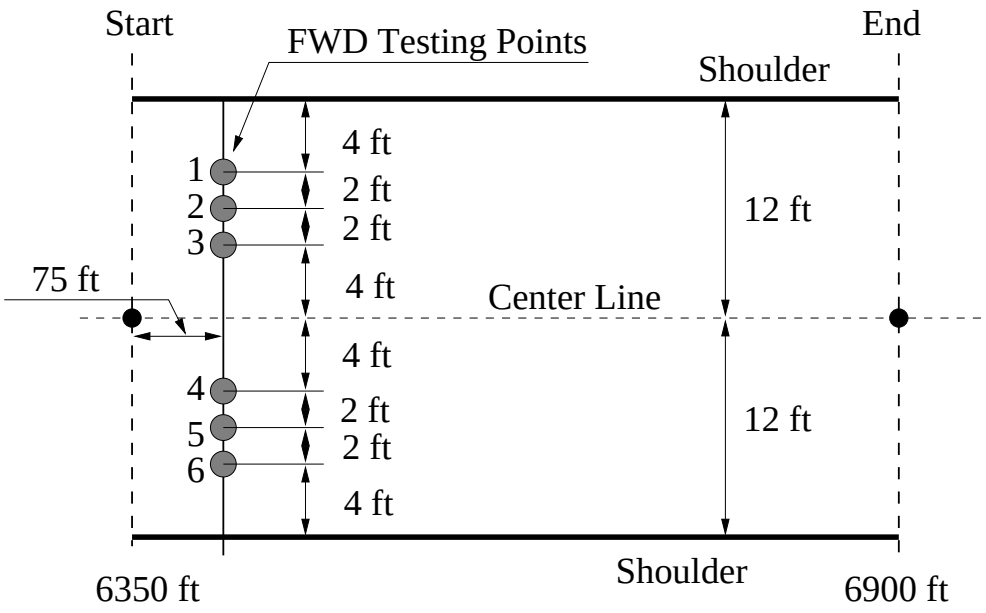

Figure 35: Testing Locations of Special FWD Test (Mn/ROAD, Low Volume Road, Test Section 33 and 33t)

FWD measurements. For the FWD tests performed at test location 3 under the load of 34 kN (Mn/ROAD Test Section 33, May 22, 2001), this FRF comparison is shown in Figs. 36 to 40. From Table 13 and these figures, it can be seen that:

- All five ANNs yield (i) mutually consistent estimates of the four output parameters (G_B, G_S, h_S and G_H), and (ii) a reasonable good agreement between the predicted and measured FRF.

- Although ANN 1 gives better results on the synthetically-generated, i.e. noise free test data than the other four ANNs (e.g. column h_S in Table 12), it does not outperform other networks on the actual FWD measurements (see Table 13).

As an illustration, Table 14 summarizes the predicted pavement properties given by ANN 5 in detail that includes six test locations (see Fig. 35) and two different impact load levels. In Table 14, the last two rows show the average value and standard deviation of the predicted values for all test location and load level combinations. From the Table,

it can be seen that ANN 5 gives consistent predictions for the FWD measurements at different test location and load level combinations. This finding holds for other four ANNs as well.

Table 13: Comparison of ANN Predictions (Average±STD), Mn/ROAD Test Section 33, 05/22/01

ANN	G_B [MPa]	G_S [MPa]	h_S [m]	G_H [MPa]
ANN 1	55 ± 12	32 ± 3	0.9 ± 0.10	90 ± 9
ANN 2	62 ± 5	33 ± 4	0.8 ± 0.05	96 ± 5
ANN 3	57 ± 13	28 ± 3	0.8 ± 0.07	102 ± 5
ANN 4	61 ± 8	31 ± 4	1.0 ± 0.05	97 ± 4
ANN 5	51 ± 10	31 ± 3	0.9 ± 0.08	91 ± 7

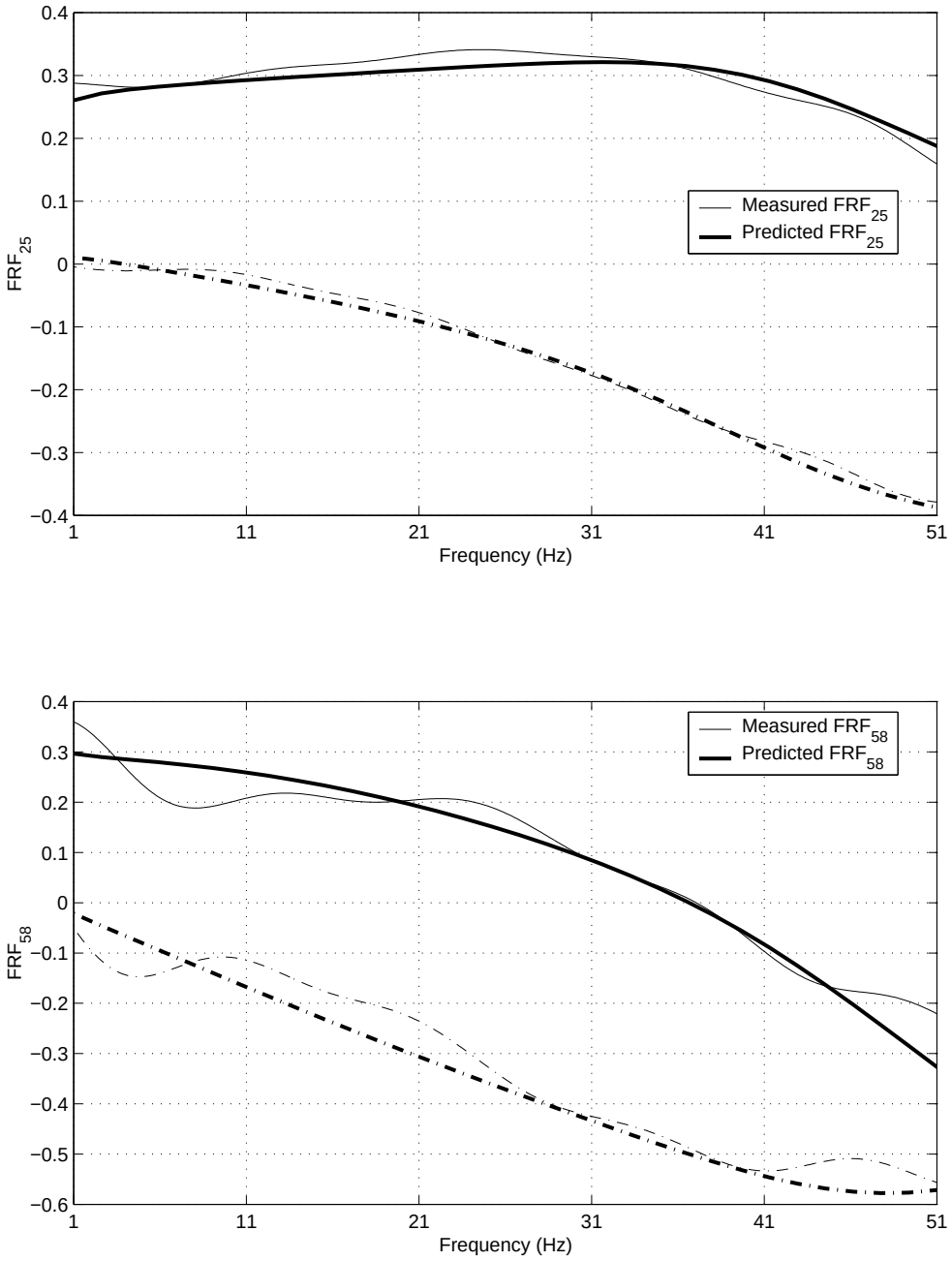

Figure 36: Prediction of ANN 1, Test Location 3, Impact Load 34 kN (Mn/ROAD Test Section 33, 05/22/01)

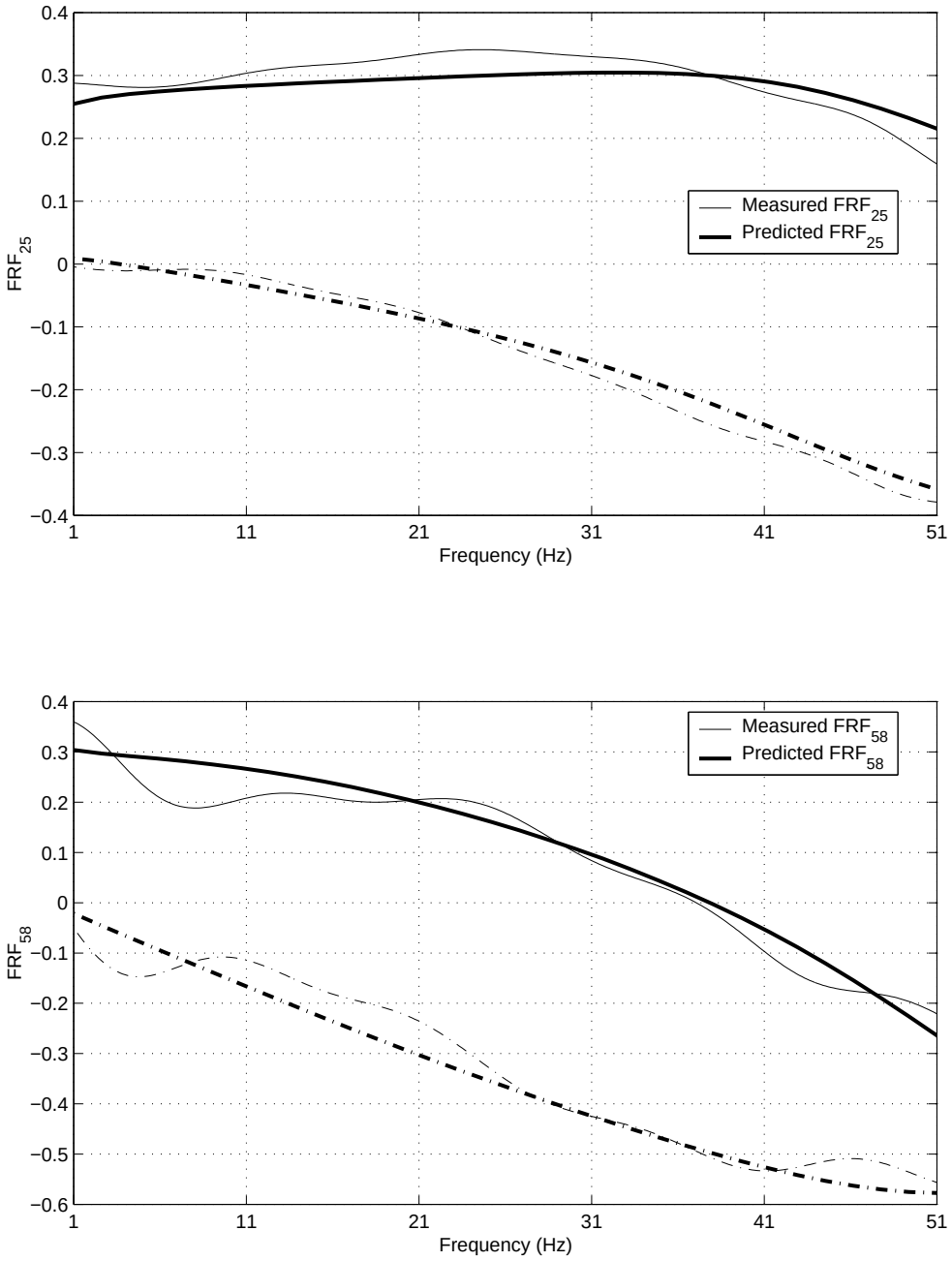

Figure 37: Prediction of ANN 2, Test Location 3, Impact Load 34 kN (Mn/ROAD Test Section 33, 05/22/01)

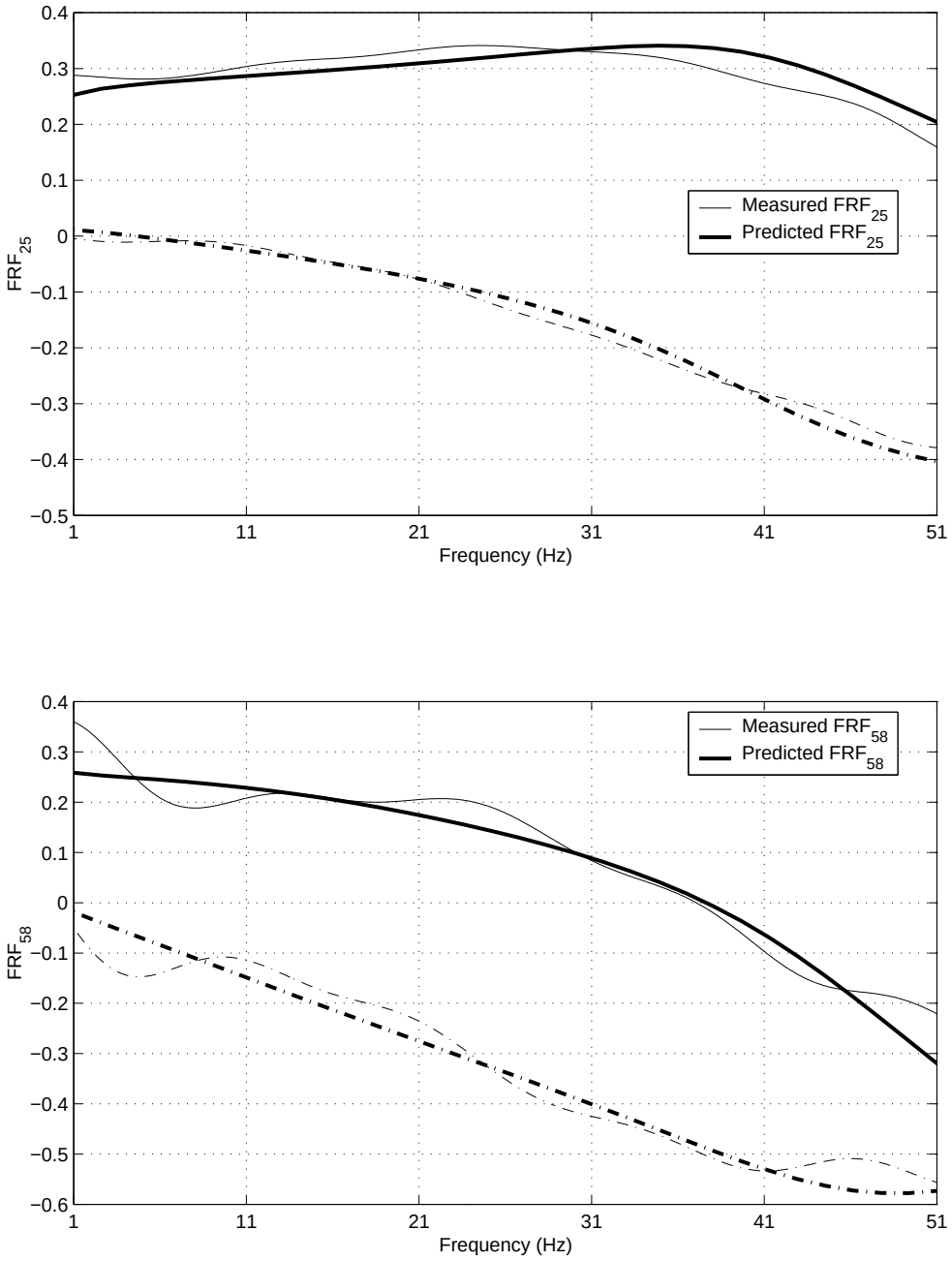

Figure 38: Prediction of ANN 3, Test Location 3, Impact Load 34 kN (Mn/ROAD Test Section 33, 05/22/01)

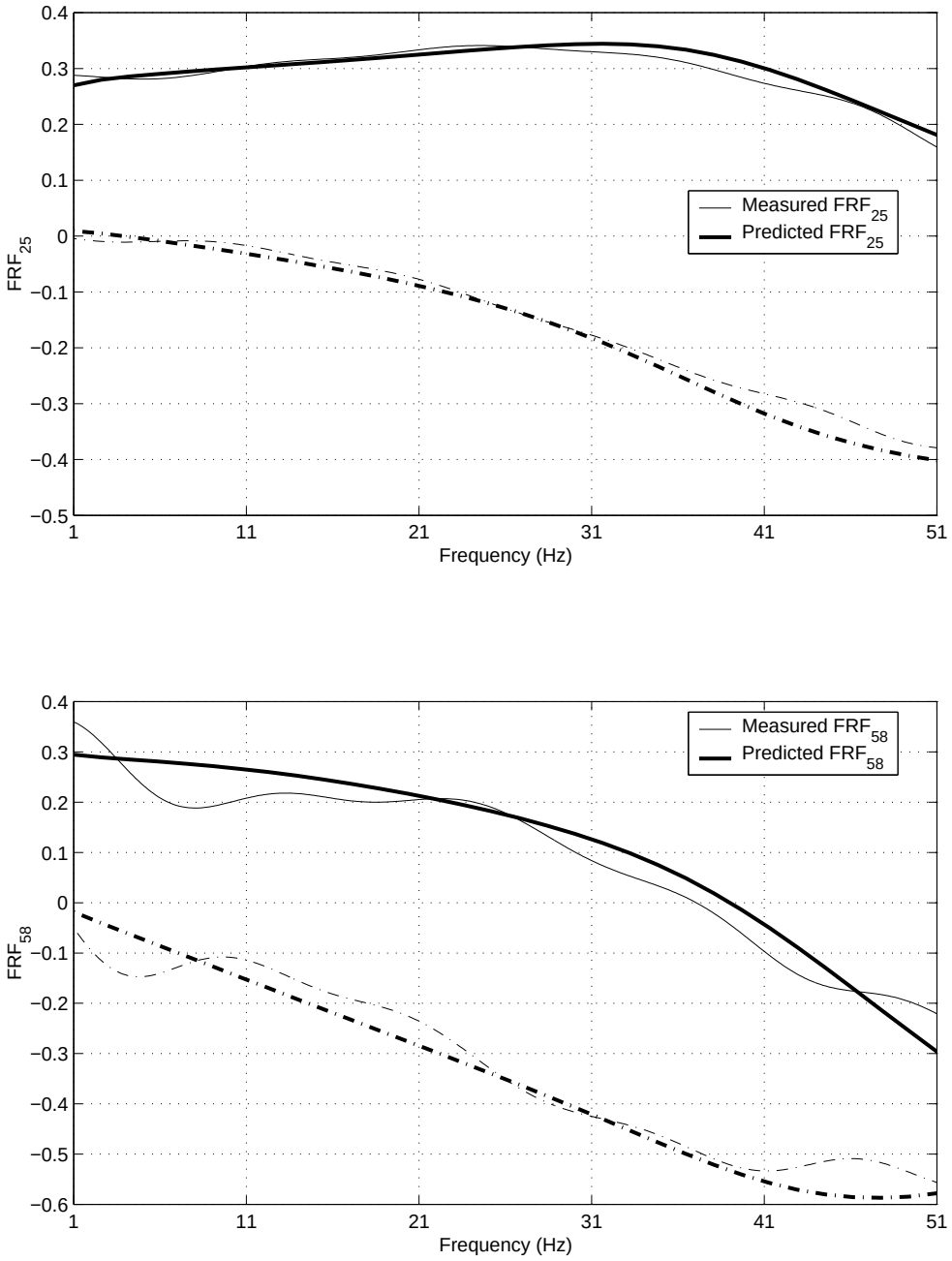

Figure 39: Prediction of ANN 4, Test Location 3, Impact Load 34 kN (Mn/ROAD Test Section 33, 05/22/01)

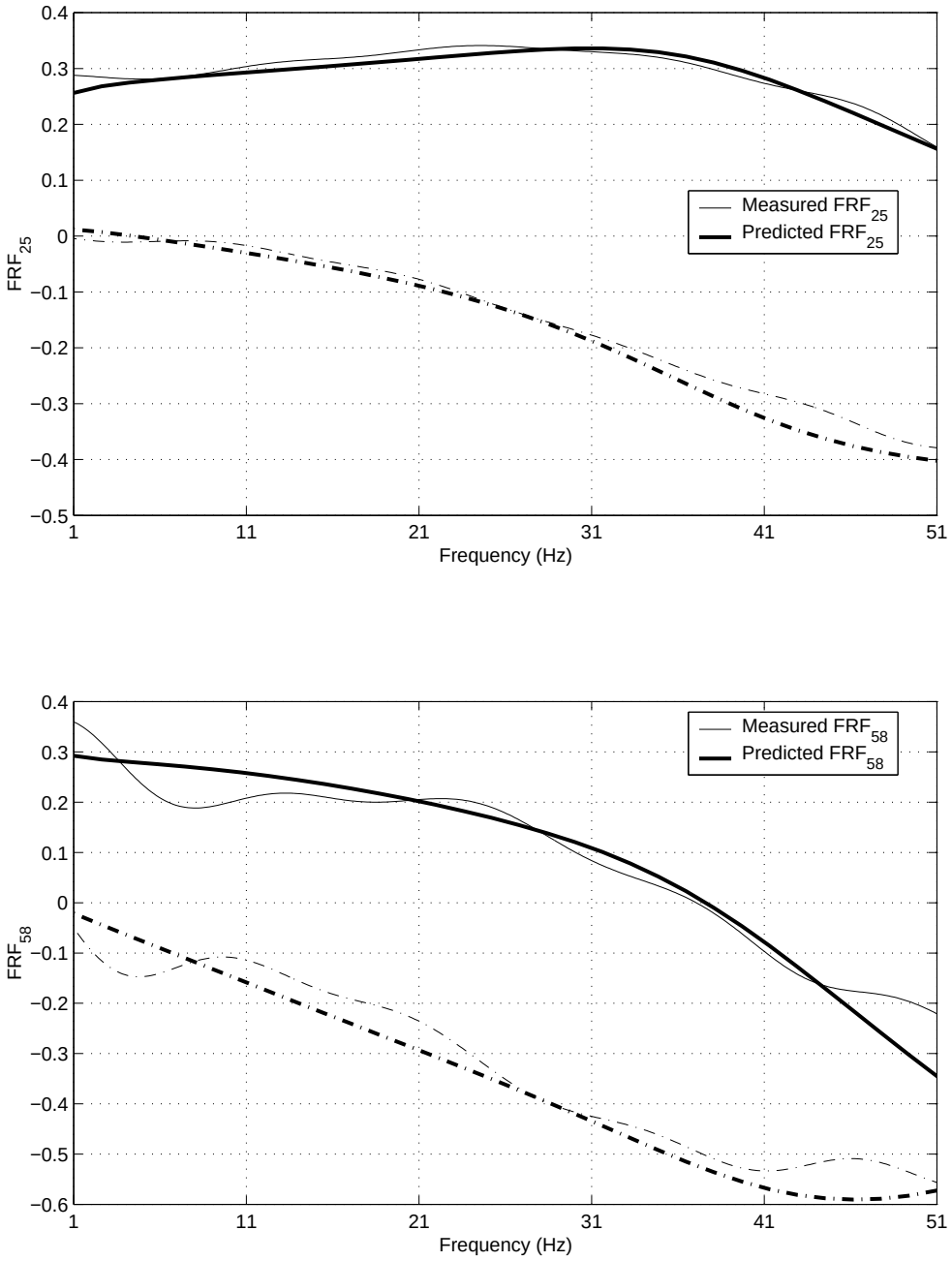

Figure 40: Prediction of ANN 5, Test Location 3, Impact Load 34 kN (Mn/ROAD Test Section 33, 05/22/01)

Table 14: Detailed Prediction Results (ANN 5), Mn/ROAD Test Section 33, 05/22/01

Point	Load Level [kN]	G_B [MPa]	G_S [MPa]	h_S [m]	G_H [MPa]
1	34	71	28	0.9	99
1	25	64	28	0.9	100
2	34	57	30	1.0	97
2	25	55	31	1.0	100
3	34	55	34	1.0	91
3	25	48	39	1.2	83
4	34	39	33	0.9	83
4	25	42	34	0.9	84
5	34	39	29	0.9	84
5	25	40	31	0.9	85
6	35	45	29	0.9	92
6	25	51	28	0.9	97
Average	N/A	51	31	0.9	91
STD	N/A	10	3	0.08	7

7.2.2 Transition Section

At the Mn/ROAD testing facility, there is a transition section located between the Test Section 33 and the west loop on the low volume road (see Fig. 34). This section will be referred to as Test Section 33t (t means transition). Fig. 41 illustrates the profile of the transition section, which means that this section *does not* have an aggregate base layer. As a result, by assuming the existence of a "fictitious" aggregate base layer with an arbitrary thickness (h_B), the ANN backcalculation model can be used to estimate the thickness of the remaining subgrade (h_S) as well as the shear modulus of the fictitious aggregate base and remaining subgrade (G_B, G_S). Because the fictitious aggregate base is in fact a part of the subgrade, the following two facts should hold if the ANN backcalculation model is capable of backcalculating the pavement properties from FWD measurements appropriately:

- The shear modulus of the fictitious aggregate base and the remaining subgrade (G_B and G_S) should be the same;
- The *sum* of the thicknesses of fictitious aggregate base and remaining subgrade layers ($h_B + h_S$) should not vary when the assumed thickness of the fictitious aggregate base (h_B) (used as an input to ANN predictions) takes on different values.

To check whether the above two requirements are satisfied, the FWD test was performed on the transition section at six different testing locations across the width of the pavement (see Fig. 35) on May 29, 2001. In the comparison, all five neural networks described in Section 6.3.1 are applied to the FWD measurements performed at the six test locations using various load levels by assuming the thickness of the fictitious aggregate base $h_B = 0.4$ m. Table 15 shows the comparison of the *average* predicted values of the shear modulus of fictitious aggregate base (G_B) and subgrade (G_S). From the column labeled $|G_B - G_S|$, it can be seen that ANN 5 gives the smallest value of $|G_B - G_S|$ among the five competing ANNs. In the table, the shear modulus is given in the form of "Average value $\pm$ Standard deviation".

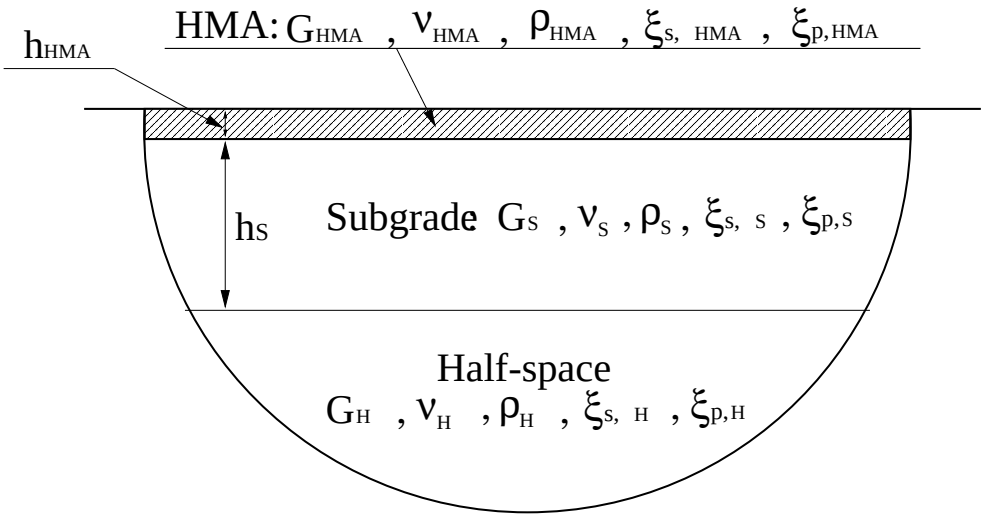

Figure 41: Pavement Profile, Test Section 33t (Transition Section)

Table 15: Comparison of G_B and G_S, Average±STD (Mn/ROAD Test Section 33t, 05/29/01), $h_B = 0.4$ m

ANN	G_B [MPa]	G_S [MPa]	$\lvert G_B - G_S \rvert$ [MPa]
ANN 1	33 ± 5	53 ± 12	20
ANN 2	51 ± 4	46 ± 5	5
ANN 3	32 ± 9	47 ± 5	15
ANN 4	47 ± 9	35 ± 3	12
ANN 5	49 ± 5	46 ± 3	3

To further investigate the performance of the neural networks studied, the assumed thickness of the fictitious aggregate base (used as an input to ANN) was varied from $h_B = 0.2$ m to $h_B = 0.7$ m with 0.1 m increments. With such input variation, the five ANNs were applied to the FWD measurements from Test Section 33t (May 29, 2001). Table 16 shows the comparison of $h_B + h_S$ as given by each ANN under different values assumed for h_B. From Table 16, it can be seen that ANN 5 yields a more consistent value of $(h_B + h_S)$ than the other four ANNs under varied input h_B.

Table 16: Comparison of $h_B + h_S$ [m], Mn/ROAD Test Section 33t, 05/29/01

h_B [m]	0.2	0.3	0.4	0.5	0.6	0.7
ANN 1	1.0 ± 0.12	1.1 ± 0.13	1.1 ± 0.14	1.1 ± 0.15	1.2 ± 0.17	1.2 ± 0.19
ANN 2	0.9 ± 0.07	1.0 ± 0.06	1.0 ± 0.06	1.1 ± 0.05	1.1 ± 0.05	1.2 ± 0.04
ANN 3	1.3 ± 0.08	1.4 ± 0.08	1.4 ± 0.10	1.5 ± 0.11	1.5 ± 0.13	1.6 ± 0.16
ANN 4	1.2 ± 0.04	1.2 ± 0.04	1.2 ± 0.04	1.2 ± 0.04	1.3 ± 0.04	1.3 ± 0.04
ANN 5	1.2 ± 0.13	1.2 ± 0.11	1.2 ± 0.11	1.2 ± 0.10	1.2 ± 0.11	1.2 ± 0.11

From the above comparisons, it appears that ANN 5 produces most consistent results among five networks with regard to the FWD measurements obtained at Section 33t. The performance of ANN 5 on the transition section is shown in detail in Table 17 where the thickness of fictitious aggregate base (h_B) is assumed to be 0.4 m. From Table 17, it can be seen that for different combinations of the testing location and load level, the selected backcalculation model (ANN 5) gives similar values for G_B and G_S and that ($h_B + h_S$) does not vary significantly between individual FWD tests. For completeness, Fig. 42 plots the comparison between the measured FRF and the predicted FRF stemming from ANN 5, from which a reasonable agreement should be observed.

Table 17: Prediction Results of ANN 5, Mn/ROAD Test Section 33t, 05/29/01 (h_B = 0.4m)

Point	Load Level [kN]	G_B [MPa]	G_S [MPa]	$h_B + h_S$ [m]	G_H [MPa]
1	34	43	42	1.0	111
1	25	46	42	1.0	113
2	34	45	45	1.1	113
2	25	50	47	1.0	109
3	34	39	46	1.3	117
3	26	49	49	1.2	108
4	35	51	46	1.3	112
4	26	55	46	1.2	113
5	35	51	48	1.3	111
5	26	52	51	1.3	108
6	35	52	42	1.1	114
6	26	53	43	1.2	115
Average	N/A	49	46	1.2	112
STD	N/A	5	3	0.11	3

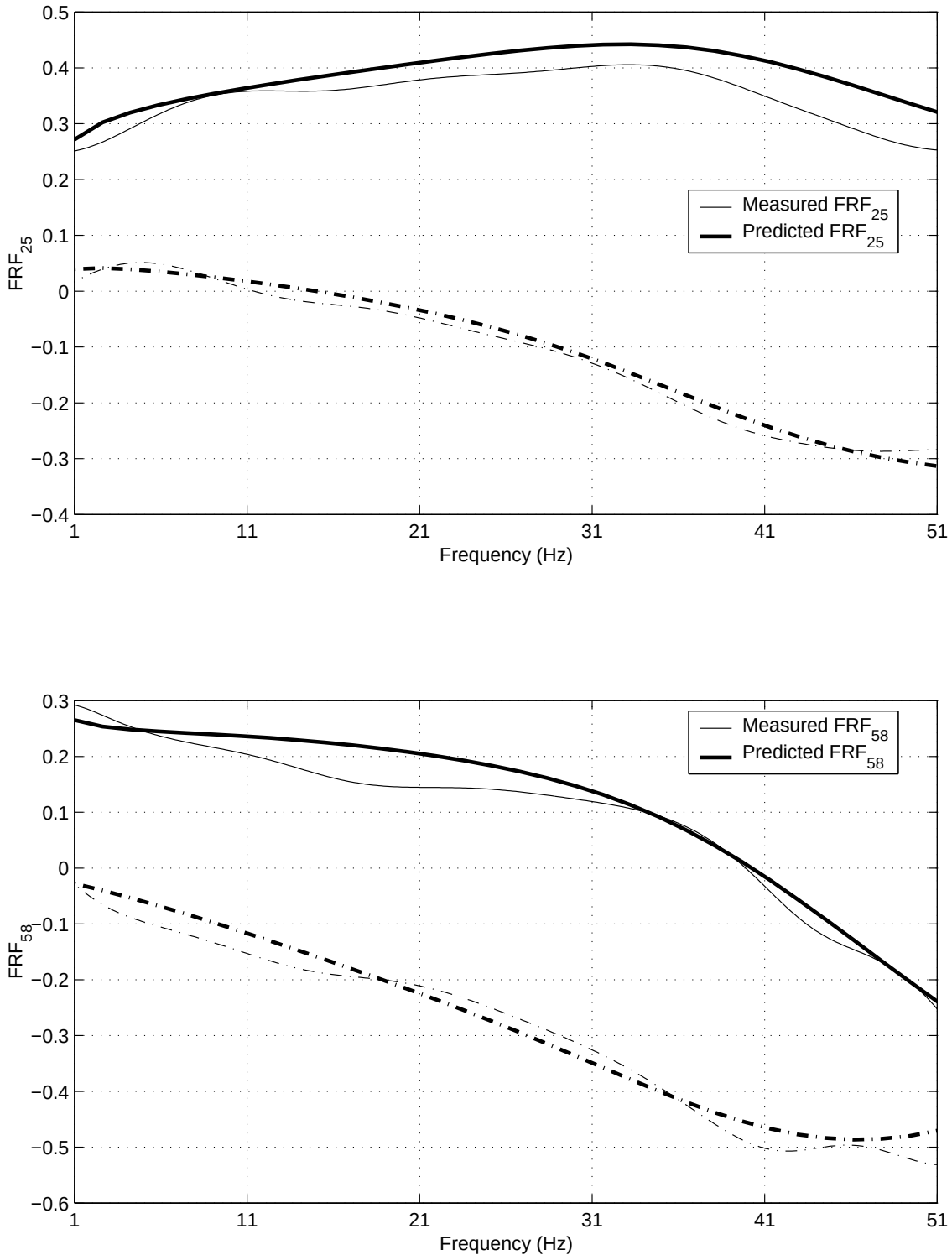

Figure 42: Prediction of ANN 5, Test Location 3, Impulse Load 34 kN (Mn/ROAD Test Section 33t, 05/22/01)

7.3 Discussion and Comments

7.3.1 Comparison with Evercalc

Evercalc ([33]) is a widely used backcalculation software based on elastostatic multi-layer analysis. Table 18 shows the comparison between the prediction results given by the proposed ANN backcalculation scheme and those given by Evercalc for Test Section 33, May 22, 2001. From the table, it can be seen that Evercalc gives a smaller value of the estimated Young's modulus for the aggregate base than for the subgrade which is contrary to the physical situation, while the ANN backcalculation scheme (the row labeled with "ANN 5" in the table) yields a reasonable result, i.e. $E_B > E_S$. It should be noted that, based on the author's experience, the depth to the stiff layer given by Evercalc is not sensitive to the stiffness of the stiff layer (E_H). Therefore, Evercalc is used as follows: First, assuming the Young's modulus of the stiff layer (i.e. half-space) to be 350 MPa, Evercalc gives the depth to the stiff layer; Second, this depth to the stiff layer is fed into Evercalc resulting in the stiffness of the stiff layer (together with other pavement parameters). Finally, the calculated depth to the stiff layer in the first step and the calculated stiffness for every layer from the second step are taken as the final results.

Table 18: ANN Prediction vs. Evercalc (Test Section 33)

Parameter	G_B [MPa]	G_S [MPa]	h_S [m]	G_H [MPa]
ANN 5	51	31	0.9	91
Evercalc	16	54	2.8	61

The comparison between the ANN and Evercalc performance is also made for Test Section 33t (May 29, 2001) as shown in Table 19, where properties specified with bold face denote the input to the corresponding backcalculation model. In the comparison, the thickness of the fictitious aggregate base takes the assumed value of 0.4 m. With reference to the iterative procedure of using Evercalc as described in the above para-

graph, the row labeled "Evercalc(1)" denotes the back-calculation results given by the first step (the Young's modulus of half-space is assumed to be 350 MPa), while the row labeled "Evercalc(2)" stands for the calculation results given by the second step. It can be seen that the difference between E_B and E_S given by ANN 5 is smaller than that given by Evercalc(1) and Evercalc(2). Actually, from the results shown in the bottom row, it can be concluded that Evercalc fails to give reasonable predictions when being applied to the FWD measurements on transition section.

Table 19: ANN Prediction vs. Evercalc (Test Section 33t)

Parameter	h_B [m]	G_B [MPa]	G_S [MPa]	h_S [m]	G_H [MPa]
ANN 5	**0.4**	49	46	0.8	112
Evercalc(1)	**0.4**	27	30	1.2	**112**
Evercalc(2)	**0.4**	10	443	**1.2**	68

7.3.2 Depth to Stiff Layer

One of the main goals of the ANN backcalculation scheme developed here is to predict the depth to the stiff layer. As a result, it is important to compare the ANN-predicted depth to the stiff layer with the one given by the construction record and other methods. For the Test Section 33 examined before, there is a construction record of the pavement profile as shown in Fig. 43. The FWD measurements from Test Section 33 that are used in the current project were performed at $64 + 25$ (see Fig. 43, $64 + 25$ means 25 feet far to the beginning of the station 64) and, from Fig. 43, the depth to the bottom of the fill soil (i.e. top of the stiff layer) is

$$d_{SL,Actual} = 0.0254 \times (6 + 6 + 48) = 1.524\text{m} \tag{75}$$

There are three ways available to the author that can be used to estimate the depth to the stiff layer: Evercalc, ANN backcalculation model developed here and the empirical formulas given by Chang et.al. [34]:

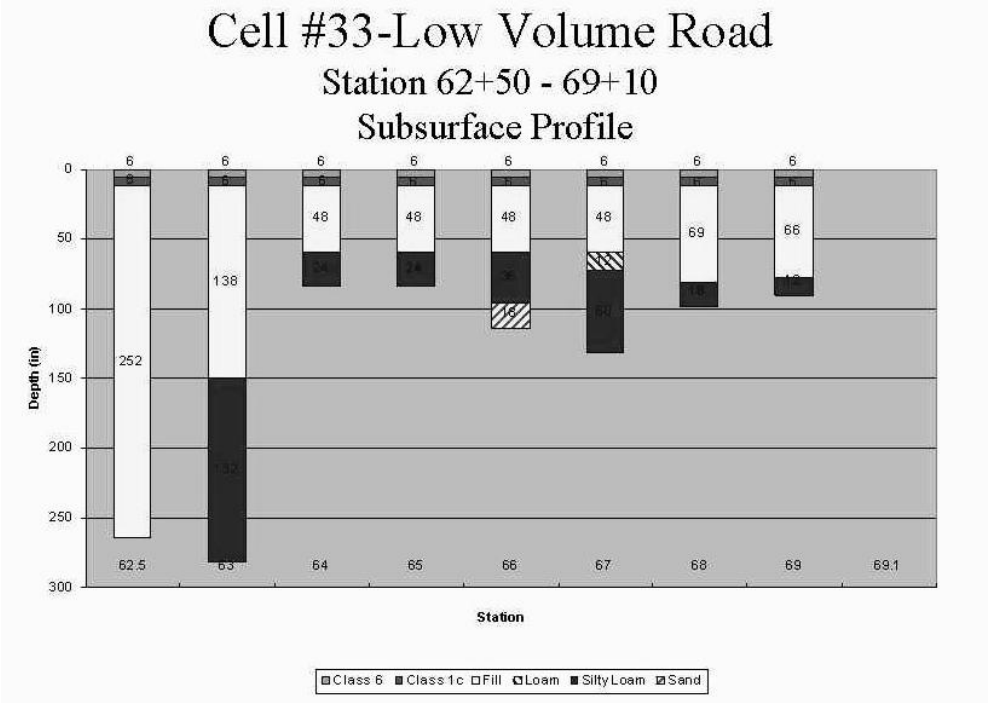

Figure 43: Actual Pavement Profile for Test Section 33 (Provided by Mn/DOT)

- **Evercalc** Based on Table 18, the depth to the stiff layer given by Evercalc is:

$$d_{SL,Evercalc} = h_{HMA} + h_B + h_S = 0.1 + 0.3 + 2.8 = 3.2\text{m} \tag{76}$$

- **ANN Backcalculation Model** On the basis of the average value of the thickness of subgrade (h_S) given in Table 14 for Test Section 33, the depth to the stiff layer (i.e. half-space) can be calculated as:

$$d_{SL,ANN\ 5} = h_{HMA} + h_B + h_S = 0.1 + 0.3 + 0.9 = 1.3\text{m} \tag{77}$$

- **Empirical Formula** According to the observation of the Chang et.al. [34], the depth to the bedrock (i.e. very stiff half-space) can be estimated based on the deflection time history obtained in the FWD Tests. As shown in Fig. 44, there are several oscillations following the main pulse, which represent the free vibrations

of the pavement system under the impulse load. Based on the observation that the period of these residual oscillations corresponds approximately to the natural period of the profile in compression, Chang proposed that the depth to the bedrock can be estimated by:

$$h = 0.25\ c_{p,S}\ T_{osc} \tag{78}$$

$$c_{p,S} = \sqrt{\frac{G_S(2-2\nu_S)}{\rho_S(1-2\nu_S)}}$$

where $c_{p,S}$ is the compressional wave speed of the subgrade, T_{osc} is the period of the residual oscillation in deflection time history (Fig. 44), G_S is the shear modulus of the subgrade, ν_S is the Poisson's ratio of the subgrade, and ρ_S is the mass density of the subgrade. Using the assumed values:

$$G_S = 70\text{MPa}, \qquad \nu_S = 0.4, \qquad \rho_S = 1865\text{kg/m}^3 \tag{79}$$

and the oscillation period estimated from Fig. 44

$$T_{osc} = 0.028\text{sec}, \tag{80}$$

the depth to the bedrock (i.e., stiff layer) is obtained:

$$d_{SL,Empirical} = 3.3\text{m} \tag{81}$$

In the above methods for estimating the depth to the stiff layer, $d_{SL,ANN\ 5}$ appears to be the most accurate when compared with $d_{SL,Actual}$. The reason for the inaccuracy of $d_{SL,Empirical}$ is that derivation of Eq. (78) assumes the stiff layer to be a bedrock, which is much stiffer than the half-space addressed in the current project, and therefore $d_{SL,Empirical}$ is an over-estimated value. Moreover, according to [33], the stiff layer in Evercalc is assumed to be 100 stiffer than the subgrade, resulting in an over-estimate of the depth to the stiff layer as well.

It should be noted that the ANN backcalculation model developed in this study has a limitation on the depth to the stiff layer that can be estimated from the range of

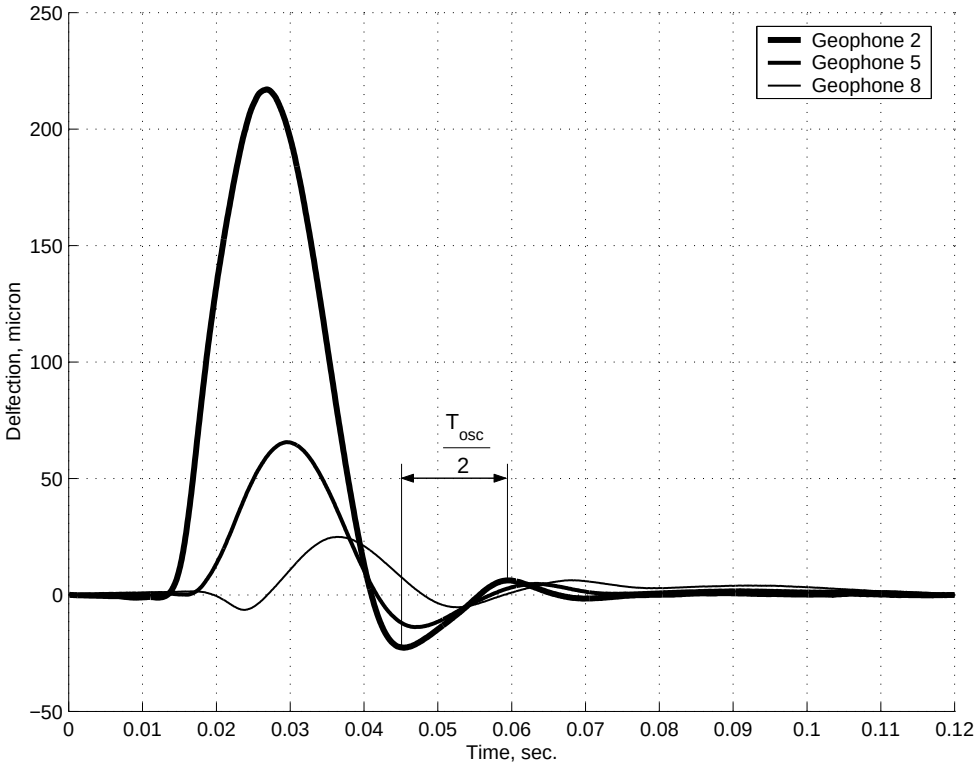

Figure 44: Deflection Time History for Test On Test Section 33 (May 22, 2001)

h_{HMA}, h_B and h_S in Table 11 as

$$d_{SL,min} = h_{HMA,min} + h_{B,min} + h_{S,min} = 0.1 + 0.1 + 0.3 = 0.5 \text{ m}$$

$$d_{SL,max} = h_{HMA,max} + h_{B,max} + h_{S,max} = 0.4 + 1.0 + 3.0 = 4.4 \text{ m}$$

On may also observe that the above depth range refers on an extreme case, and the actual range of the depth to the stiff layer recognizable by the ANN backcalculation model developed here is determined by:

$$[h_{HMA} + h_B + 0.3 \text{ m}, h_{HMA} + h_B + 3.0 \text{ m}] \tag{82}$$

where h_{HMA} and h_B are determined by the specific pavement profile.

7.3.3 Baseline Correction

Because of the experimental noise embedded in FWD measurements, the baseline correction must be applied to the deflection time histories before they can be used to estimate the FRF's as examined in Section 4.3.3. As shown in Fig. 19, there is a visible difference between the experimental FRF's at lower frequencies depending on the baseline correction method used. Table 20 shows the effect of baseline correction on the backcalculation results. It can be seen that, although the only significant difference of FRF's occurs below 10 Hz, there is a substantial effect of this difference on the predicted values of G_B and G_H. It is anticipated that further investigation and development of an optimal baseline correction method would improve the prediction quality of the backcalculation model developed. In this study, the linear baseline correction is recommended and set as the default baseline correction method.

Table 20: Effect of Baseline Correction on Prediction Results (ANN 5, Mn/ROAD Test Section 33, 05/22/01)

Method	G_B [MPa]	G_S [MPa]	h_S [m]	G_H [MPa]
Linear Correction	51 ± 10	31 ± 3	0.9 ± 0.08	91 ± 7
Quadratic Correction	35 ± 8	26 ± 2	0.9 ± 0.14	119 ± 5
Square Root Correction	62 ± 12	33 ± 3	1.0 ± 0.10	76 ± 9

8 Summaries and Conclusions

In this investigation, a novel backcalculation scheme is developed that is capable of reliably estimating the location and properties of the permanent or seasonal stiff layer (underlying asphalt pavements) from faling weight deflectometer records. Performance of the method is validated through field measurements.

The backcalculation technique proposed is based on the Artificial Neural Network (ANN) approach as a *pattern recognition tool* and a three-dimensional multilayer viscoelastodynamic model as a *predictive tool*. The key features of the proposed backcalculation technique are summarized below.

1. **Resolution Limits of Seismic Waves:** Recognizing the inherent inability of low-frequency seismic waves (such as those generated by the FWD impulse loading) to resolve thin layers, material properties of the asphalt layer are assumed to be *known* and are therefore used as an *input* to the ANN-based back-analysis.

2. **Material Behavior of Asphalt Concrete:** The Hot Mix Asphalt (HMA) layer is modeled as a viscoelastic material using the power law model. For a specific type of hot-mix asphalt, the power-law parameters are estimated via empirical formulas as a function of the mix characteristics and daily temperature.

3. **Representation of Field Measurements:** The FWD measurements are interpreted in terms of the pavement's frequency response functions (FRF's) which are shown to be an effective tool to characterize any linear (pavement) system in the frequency domain and thus expose its resonance and stiffness characteristics.

4. **Backcalculation Strategy:** The inputs to the ANN-based backcalculation of pavement stiffness characteristics are (i) the frequency response functions representing the FWD measurements, (ii) the coefficients of the power law model synthesizing the asphalt properties, and (iii) the thicknesses of the asphalt and aggregate base layers. The outputs of the backcalculation scheme are (i) the shear

moduli of aggregate base, subgrade, and the stiff layer, as well as (ii) the thickness of the subgrade layer.

5. **Integral Sampling Scheme:** Like any experimental data, it is inevitable that the FWD measurements are noise contaminated, which in turn influences the estimated frequency response functions. To render the backcalculation model robust in the presence of noisy measurements, the experimental frequency response functions (synthesizing the pavement response to dynamic loading) are interpreted in terms of their segmental areas (see Section 4.1.5). On applying the backcalculation model to field FWD data, this approach has been shown to be very effective.

6. **Lateral Wave Reflections in Pavement Testing:** Due to the finite width of the pavement section, wave reflections from lateral pavement edges are found to influence the FWD measurements significantly, a phenomenon that is not accounted for by the predictive (i.e. forward) model employed. To minimize the effect of side reflections, the useful frequency range for the frequency response functions that are used as an input to the back-analysis is taken to be from 1 to 51 Hz, thus avoiding the pronounced resonance peak due to lateral wave reflections which is typically located between 50 and 70 Hz. To further alleviate the problem associated with lateral reflections in FWD testing, it is recommended that the FWD test be performed near the center of the pavement (i.e. away from pavement edges).

7. **ANN Architecture:** Because a backpropagation neural network with two hidden layers is capable of representing the same input/output relationship with a reduced number of neurons when compared to the neural network with a single hidden layer, the neural network with two hidden layers is used in this study. However, there are no generally accepted rules regarding how to determine the number of neurons in each hidden layer for a given problem. In the current project, the preferred "rule of thumb" is that the number of neurons in the first hidden layer (next to the input layer) is twice the number of neurons in the second hidden layer (located next to the output layer).

8. **Sequential Approach:** As suggested by Kim et.al. [8], two neural networks were used sequentially in the backcalculation process. For a given input vector, the first network predicts the thickness of the subgrade, which is related to the depth to the stiff layer. On employing the estimated thickness of the subgrade to enrich the original input vector, the second network predicts the stiffnesses of the base, subgrade, and the stiff layer. This approach has been shown to have a superior performance relative to the direct approach, in which only one neural network is used.

9. **Noise Injection:** Neural network training using the noise injection technique is necessary for ANN to have a good performance on the field data. On the basis of developments in this project, it was found that the performance of noise-free neural networks on the field data is inferior to that of the ANN's trained with noise injection. For instance, through the comparison of alternative backcalculation algorithms using field measurements, it was found that the ANN's with a relative noise level of 20% (applied to input data) perform significantly better than the networks trained with a relatively low noise level of 3%. Moreover, the neural networks with the noise level tailored to each input component were found to perform better than the networks with the uniform noise level for all input parameters.

10. **Measurement Quality:** As examined before, the coherence function is a convenient criterion to judge whether the FWD measurements have an acceptably low level of noise. However, the coherence criterion is only a measure of the relative noise in experimental observations, and not a check of the measurement accuracy. Based on the parametric study using field measurements, another criterion for estimating the quality of the FWD measurements has been established that is based on the deflection time history measured by the center geophone (see Section 4.3.2). The second criterion has been effectively applied to all field measurements available to the author.

11. **Computational Efficiency:** Compared with the conventional minimization

methods such as gradient-based techniques, the use of artificial neural network in the backcalculation method reduces the computation time by more than two orders of magnitude.

To further improve the performance of the backcalculation model developed, the following recommendations are suggested:

- The mechanical properties of the asphalt layer are characterized by the power law model. Currently, the coefficients in power law model are obtained by fitting the empirical equation describing the Young's modulus of hot-mix asphalt (HMA) [17], whose inputs are the HMA aggregate coefficients, HMA temperature, and the viscosity of asphalt binder. A method that can give a more accurate estimate of the power law coefficients would certainly be beneficial.

- The noise in the FWD measurement requires the application of baseline correction to the deflection time histories (obtained by integrating the measured velocity records) before their use as a basis for the back-analysis. In this study, a significant sensitivity of the FRF estimates to the choice of baseline correction method is observed, resulting in an observable variation of the estimated pavement properties. Therefore, a more reliable baseline correction method would improve the backcalculation performance. At the same time, according to the definition of the frequency response function and Fourier transform, the FRF's of interest in this study can be directly obtained from either deflection, velocity, or acceleration time histories. As a result, a direct use of the original, i.e. measured velocity records (which are currently unavailable through the FWD setup) would significantly improve the prediction capabilities of the backcalculation scheme proposed.

- In order to detect seasonal or permanent stiff layer at larger depths (> 3 m), the use of geophones at distances larger than those available in the current FWD setup would be necessary. In that case, however, the neural network would have to be retrained using an appropriate training data set.

References

[1] Khazanovich L. and Roesler J., DIPLOBACK: Neural-Network-Based Backcalculation Program for Composite Pavements, *Transportation Research Record*, No. 1570, page 143∼150, 1997.

[2] Meier R. W. and Rix G. J., Backcalculation of Flexible Pavement Moduli From Dynamic Deflection Basins Using Artificial Neural Network, *Transportation Research Record*, No 1473, page 72 ∼ 81, 1995.

[3] Magnuson A. H., Lytton R. L., and Briggs R. C., Comparison of Computer Predictions and Field Data for Dynamic Analysis of Falling Weight Deflectometer Data, *Transportation Research Record*, issue 1293, page 61∼71, 1991.

[4] Guzina B. B., Pak R. Y. S., On the Analysis of Wave Motions In A Multi-Layered Solid, *Q. Jl Mech. Appl. Math.* page 13∼37, 54 (1), 2001.

[5] Harichandran R. S., Mahmood T., Raab A. R. and Baladi G. Y., Modified Newton Algorithm for Backcalculation of Pavement Layer Properties, *Transportation Research Record*, page 15∼22, No. 1384, 1993.

[6] Hossain A. S. M. Mustaque and Zaniewski J. P., Detection and Determination of Depth of Rigid Bottom in Backcalculation of Layer Moduli from Falling Weight Deflectometer Data, *Transportation Research Record*, page 124∼135, No. 1293, 1991.

[7] Fwa T. F., Tan C. Y. and Chan W. T., Backcalculation Analysis of Pavement-Layer Moduli Using Genetic Algorithms, *Transportation Research Record*, page 134∼142, No. 1570, 1997.

[8] Kim Y. and Kim Y. R., Prediction of Layer Moduli from Falling Weight Deflectometer and Surface Wave Measurements Using Artificial Neural Network, *Transportation Research Record*, Issue 1639, page 53∼61, 1998.

[9] Williams T. P. and Gucunski N., Neural Networks For Backcalculation Of Moduli From SASW Test, *Journal of Computing in Civil Engineering*, page 1∼8, 1995.

[10] Farr J. B., Seismic Wave Attenuation and Rock Properties, page 302∼320, *Site Characterization & Exploration*, Published by American Society of Civil Engineers, ISBN 0-87262-186-3, 1979.

[11] Guzina B. B. and Nintcheu F. S., Axial vibration of a padded annulus on a semi-infinite viscoelastic solid, *Quart. J. Mech. Appl. Math*, 2001.

[12] Findley W. N., Lai J. S. and Onaran K., *Creep and Relaxation of Nonlinear Viscoelastic Materials: with An Introduction to Linear Viscoelasticity*, Dover Publications, INC., New York, ISBN: 0-486-66016-8, 1989.

[13] Zhang W., Drescher A., Newcomb D. E., Viscoelastic Analysis of Diameteral Compression of Asphalt Concrete, *Journal of Engineering Mechanics*, page 596∼603, June 1997.

[14] Ewing W., Jardetsky W. and Press F., *Elastic Waves in Layered Media*, McGraw-Hill Book Co., New York, 1957.

[15] Sneddon I. N., *Fourier Transforms*, McGRAW-Hill Book Company, Inc. 1951.

[16] Hopman P. C., VEROAD: A Viscoelastic Multilayer Computer Program, *Transportation Research Record*, Issue 1539, page 72∼80, 1996.

[17] Witczak M. W. and Fonseca O. A., Revised Predictive Model for Dynamic (Complex) Modulus of Asphalt Mixtures, *Transportation Research Record*, issue 1540, 1996.

[18] Christensen R. M., *Theory of Viscoelasticity*, Academic Press, New York, 1971.

[19] Nazarian S. and Stokoe K.H. II, Nondestructive Evaluation of Pavements by Surface Wave Method, *ASTM special technical publication.*, 1026, eds. Bush A. J. III and Baladi G. Y., page 119 $\sim$ 137, 1989.

[20] Bendat J. S. and Piersol A. G., *Random Data Analysis and Measurement Procedures, Third Edition*, Published by Wiley-Interscience, 2000.

[21] Russell, S. J. and Norvig, P., *Artificial Intelligence: A Modern Approach*, Prentice Hall PTR, December 1994, ISBN: 0131038052

[22] Mathworks Inc., The Matlab user's manual for MATLAB Version 5.3.0.10183 (R11) and Neural Network Toolbox Version 3.0.1 (R11) 01-Jul-1998.

[23] This is a web page that includes a rather comprehensive introduction various aspects of artificial neural networks. The information used here comes from ftp://ftp.sas.com/pub/neural/FAQ2.html.

[24] Rumelhart, D. E., Hinton, G. E., and Williams R. J., Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, and the PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Volume 1: Foundations, page 318 $\sim$ 362. Cambridge, MA, MIT Press, 1986. Sixth printing at 1987.

[25] Plaut, D., Nowlan, S. and Hinton, G. E., Experiments on learning by back propagation. Technical Report CMU-CS-86-126, Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1986.

[26] Fahlman, S. E., An Empirical Study of Learning Speed in Back-Propagation Networks, Technical report at Carnegie Mellon University, CMU-CS-88-162, ftp://ftp.cs.cmu.edu/afs/cs/project/connect/tr/qp-tr.ps.Z, September, 1988.

[27] Hagan M. T. and Menhaj M. B., Training Feedforward Networks with the Marquardt Algorithm, *IEEE Transactions on Neural Networks*, Vol 5, No. 6, November 1994.

[28] Caruana R., Lawrence S. and Giles L., Over fitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping, *Neural Information Processing*, Denver, Colorado, November 28$\sim$30, 2000.

[29] Bartlett P. L., For valid generalization the size of the weights is more important than the size of the network, *Advances in Neural Information Processing Systems*, volume 9, page 134. The MIT Press, 1997.

[30] Duda R. O., Hart P. E. and Stork D. G. *Pattern Classification, second edition*, Wiley-Interscience Publication, 2000.

[31] Matsuoka K., Noise Injection into inputs in back-propagation learning, *IEEE Transactions Systems, Man. and Cybernetics*, vol. 22, No. 3, page 436∼440, 1992.

[32] Grandvalet Y. and Canu S., Comments on "Noise Injection into Inputs in Back Propagation Learning", *IEEE Transactions on System, Man and Cybernetics*, Vol. 25, No. 4, page 678∼681, April 1995.

[33] Washington State Department of Transportation, *WSDOT Pavement Guide*, Volume 2 and 3, Seattle, Washington, 1995.

[34] Chang D. W., Kang Y. V., Roesset J. M. and Stokoe II K. H., Effect of Depth to Bedrock on Deflection Basins Obtained with Dynaflect and Falling Weight Deflectometer Tests, *Transportation Research Record*, 1355, page 8∼16, 1992.

[35] Meier R. W. and Rix G. J., Backcalculation of Flexible Pavement Moduli From Dynamic Deflection Basins Using Artificial Neural Networks, *Transportation Research Record* 1473, page 72∼81, 1995.