Identifiability and Parameter Estimation

in Rail Vehicle Dynamics

by

Bradley M. Coffey

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Mechanical Engineering

APPROVED:

Dr. Robert H. Fries, Chairman

Dr. Charles F. Reinholtz

Dr. Robert G. Leonard

May 24, 1988

Blacksburg, Virginia

LD
5655
V855
1988
C633
c.2

Identifiability and Parameter Estimation

in Rail Vehicle Dynamics

by

Bradley M. Coffey

Dr. Robert H. Fries, Chairman

Mechanical Engineering

(ABSTRACT)

Rail vehicle designers and analysts can benefit from the results of vehicle parameter estimation. Using this technique, they can determine the effects of suspension design decisions, and they can reduce the amount of on-track testing required to qualify new designs for service.

This work addresses two major issues: the determination of parameter identifiability and the estimation of rail vehicle parameters from laboratory tests. Usually, the identifiability issue should be addressed first since identifiability determines the number of independent parameters that can be estimated.

The general issues of identifiability and parameter estimation are discussed. Two identifiability tests are explored in depth, as is a Bayesian least-squares parameter estimation method. Laboratory tests from a lightweight intermodal rail vehicle with single-axle trucks provided the data for the parameter estimation. The test setup and a simple vehicle mathematical model provided the structure for the identifiability determination.

This work shows that identifiability and estimation issues closely interact. Even if a system is not identifiable, the Bayesian estimation method can return results. Thus, the Bayesian method can instill false confidence in the validity of the estimation results.

Estimation of experimental data with a linear model provided values within one percent for the mass and damped natural frequency, and ten percent for the peak amplitude. Excellent agreement with the experimental data was obtained for frequencies above the resonant peak and for very low frequencies. Error at frequencies slightly below the resonant peak, however, indicated the vehicle contained significant nonlinearities. To achieve closer agreement between model response and test response at these frequencies, a nonlinear vehicle model is needed.

Acknowledgements

I would like to thank Dr. Robert H. Fries for providing encouragement, guidance, and support while serving as my thesis advisor. His many hours of selfless consultation are greatly appreciated. I also thank Dr. Robert G. Leonard and Dr. Charles F. Reinholtz for serving on my committee.

I am grateful for the funding from both the Virginia Center for Innovative Technology and the Mechanical Engineering department to perform this research. I thank Nicholas G. Wilson, Firdausi D. Irani, John A. Elkins, and others at the Transportation Test Center of the U.S. Department of Transportation for providing the experimental data. In addition, Burt Avery of User Services was invaluable in the translation of many data tapes. Thanks also go to Steve Wampler and Bob Arenburg for their help with computer difficulties.

I am grateful to the 13 office mates I have had in the past two years. They helped make this time, especially the late nights, much more enjoyable. I sincerely thank my many international friends for giving me a broader view of the world we all share.

Finally, I thank my parents for their unconditional love and support in my many different endeavors. Through reading endless books and explaining countless happenings to me as a child, they instilled in me a love of learning and teaching that continues today.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

Introduction and Literature Review

1.1 Introduction

The number of new designs for freight cars and trucks increased substantially in the past several years. To accommodate this increase, both the Association of American Railroads (AAR) and the Federal Railway Administration (FRA) introduced procedures to expedite the analysis and evaluation of these designs. Specifically, the FRA sponsored a project to determine efficient procedures for the analysis and testing of new designs of lightweight cars and trucks (Irani, et al., 1986).

The new rail vehicle designs result from an industry demand for equipment to carry intermodal traffic. Intermodal loads, usually one or two truck trailers or containers, are much lighter than normal loads such as on a hopper car or boxcar. The low vehicle loadings enable the design of single-axle suspension systems compared to the normal three-piece truck and suspension. Also, to maximize the

load-to-tare ratio and minimize the energy consumption, the vehicle bodies are being designed with lightweight structures. The lightweight structures may also include materials previously untested in the railroad environment. Figure 1 shows a sketch of a typical intermodal vehicle loaded with a road trailer.

With many unknowns about the vehicles, the FRA developed techniques to evaluate their safety aspects. Also, to fairly compare the different designs and to shorten the certification process, standardized laboratory testing was required. A primary task of the FRA's program was to develop partially validated mathematical computer models from the laboratory testing. The mathematical models could then be used to test the vehicle response to a wide variety of rail inputs and track classes. This computer testing greatly reduces the amount of on-track testing of the vehicle and expedites the evaluation process. However, if the parameter values in the mathematical model are inaccurate or if the model is oversimplified, the model is invalid.

This thesis provides a method for determining a linear mathematical model of a rail vehicle from controlled laboratory experiments. Two basic tasks determine the development of a good mathematical model (Junkins, 1978). First, a mathematical model is formulated consistent to physical laws governing the system's behavior, measurements available, and accuracy desired. These mathematical models usually contain imprecisely known parameter values. Second, the best estimates of the poorly known parameters are determined to form an optimal estimate of the system's actual behavior.

Unfortunately, ***parameter*** is "wordnapped into having meanings of consideration, factor, variable, influence, interaction, amount, measurement, ... , cause, effect, modification, alteration, computation,..." (Kilpatrick, 1984). Also, a parameter cannot be defined only as a constant in a mathematical equation as in

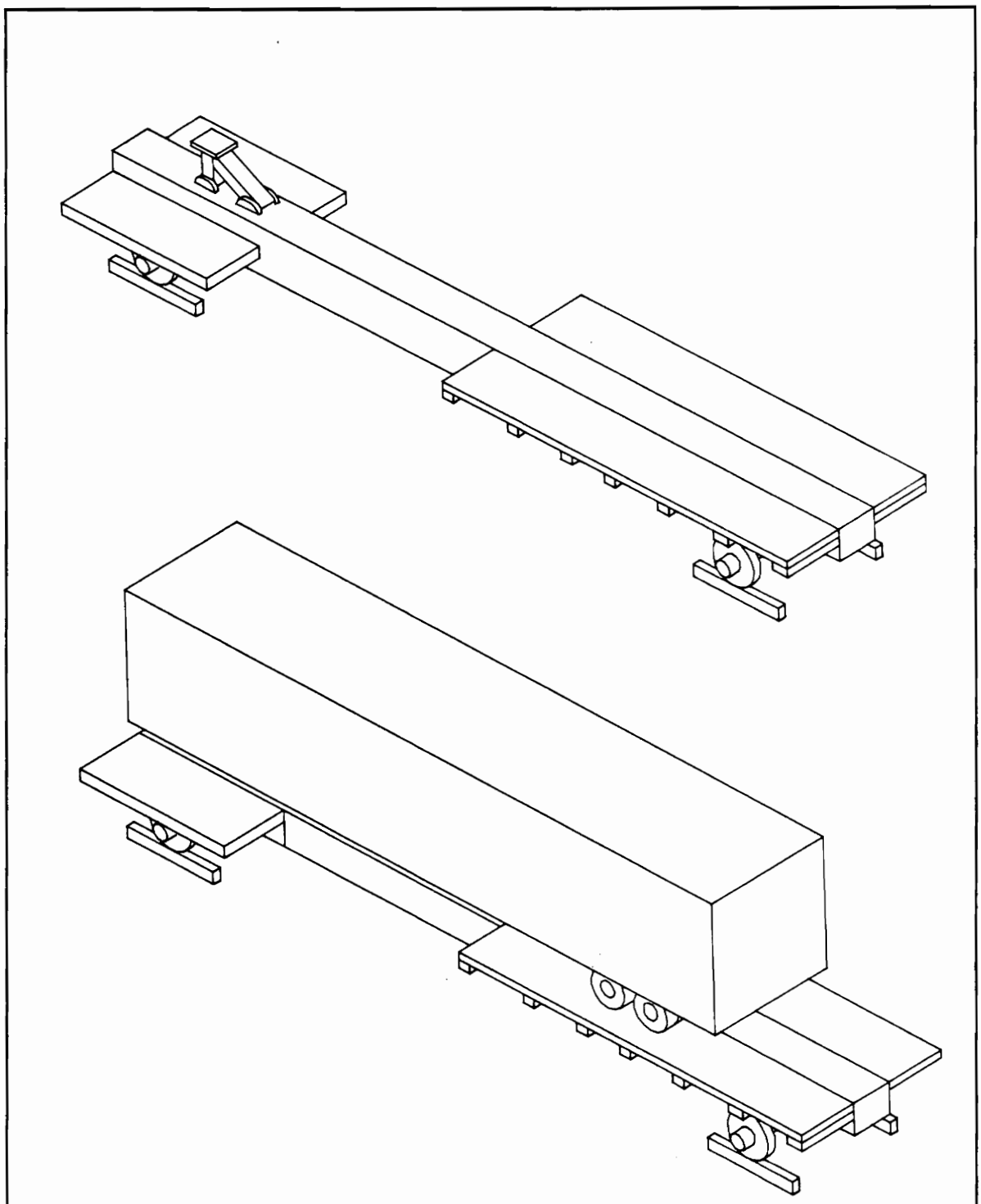


Figure 1. Typical intermodal vehicle

Beck and Arnold (1977). This definition fails when nonlinear phenomena such as Coulomb friction are modeled. In this thesis, a parameter represents a physical property of a system such as distance, mass, inertia, damping, or stiffness. These parameters appear in mathematical expressions to describe the behavior of the physical system.

Confusion also exists with the terms ***system identification*** and ***parameter estimation.*** System identification is the first task in mathematical modeling mentioned above by Junkins. Parameter estimation encompasses the second task. Combinations of the four words such as "system estimation" and "parameter identification" appear throughout the literature. Usually, to determine which task an author is pursuing, the key words are ***system*** and ***parameter***, not ***identification*** and ***estimation***.

Once the mathematical structure is defined, a natural question arises: is it possible to determine some or all of the unknown parameters in the model from the data? This is the identifiability problem. Identifiability is generally independent of the parameter estimation method used. However, both model structure and parameter values affect identifiability. Identifiability for almost all parameter values is known as structural or global identifiability. Identifiability for specific value of the parameters is called local identifiability.

Unfortunately, identifiability problems wear many different masks and are almost invariably hidden within the parameter estimation task of mathematical modeling. For example, researchers have suggested that "parameters obtained are not reasonable", "a model may be adjusted to match certain data without actually improving the the model", and "parameters may be changed without affecting the performance index". These difficulties, though all occurring in the estimation task,

represent identifiability problems with the model structure. A demonstration of this problem and analytical methods to predict it are presented in this thesis.

The estimation method used in this work is the Bayesian least-squares method. This method is widely used in rail vehicle parameter estimation. Hasselman and Johnson (1979), Gilan (1981), and Fries (1983) used this estimation method successfully. Both Gilan and Fries carefully investigated the costs and benefits of various parameter estimation methods before using the Bayesian least-squares method.

The output error formulation of the Bayesian estimation is implemented in the frequency domain for this work. The output error method minimizes a quadratic performance index containing the differences between the measured and expected outputs. The frequency domain implementation allows easier interpretation of the results and easier matching of the experiment and model. Once again, these methods prevail in the rail vehicle parameter estimation literature.

The vehicle investigated is a single-axle car with a European International Union (UIC) suspension system. This vehicle is typical of the new designs of lightweight cars. Only linear suspensions are considered. The mathematical modeling in this thesis describes only the rigid, vertical dynamics of the lightweight car. The modes investigated are bounce, pitch, and roll. Flexible mode vibration does not affect these three modes. The lateral and vertical modes are assumed uncoupled.

This thesis provides a method for determining linear mathematical models and applies to fields other than rail vehicle modeling. Biology, econometrics, and ecology represent areas outside engineering using these principles. Identifiability is one difficulty encountered in modeling real systems, regardless of the field.

Therefore, conclusions from this work may lessen difficulties in estimating parameters of dynamic systems.

1.2 Literature Review

1.2.1 Rail Vehicle Modeling and Estimation Literature

Several works enable engineers unfamiliar with rail vehicle dynamics to gain a broad understanding of the principles and problems faced today in prediction and analysis of rail vehicle dynamics. Law and Cooperrider (1974) provided a classic survey on railway vehicle dynamics and the problems researchers faced at that time. In a compilation of research papers, Moyar et al., (1978) reinforced the observations of Law and Cooperrider. A textbook by Garg and Dukkipati (1984) contains most of the information needed to develop reasonable mathematical models of modern rail vehicles. This text also includes references to recent research in rail vehicle dynamics.

Three papers investigate characteristics of lightweight cars very similar to the vehicle investigated in this work. Wormley and Tombers (1983) investigated the basic stability characteristics of the lightweight intermodal car. The methods in Wormley and Tombers' paper resulted from more general research by Hedrick et al., (1978,1979). Irani et al., (1986) provided much of the motivation for this thesis. They investigated developments in the testing and analysis of new vehicle designs. Difficulties they met in estimating the vehicle parameters led to the investigation of

identifiability in this thesis. Finally, Paul et al., (1983) considered aerodynamic properties of lightweight intermodal cars.

The intermodal vehicle was tested at the Transportation Test Center in Pueblo, Colorado. The facilities at the test center are described in papers by Inskeep and Roberts (1982) and Irani et al., (1986). Inskeep and Roberts explained the mission of the Rail Dynamics Laboratory (RDL) and provided the basic capabilities of the Roll Dynamics Unit (RDU) and the Vibration Test Unit (VTU). The intermodal vehicle was tested only on the VTU. More detailed information on the facilities at the Transportation Test Center is included in their brochures and manuals (Transportation Test Center).

Numerous authors identify parameters of vehicle dynamics models. In a survey paper, Kallenbach (1987) introduced identification techniques for complex vehicle models. Kallenbach developed several estimation approaches and identified the parameters of vertical automobile dynamics.

Specifically for rail vehicle dynamics, Hasselman and Johnson (1979) used Bayesian statistical parameter estimation to determine the rock and roll dynamics of a four degree of freedom 100-Ton hopper car. Hasselman and Johnson's work is the first known application of Bayesian statistical parameter estimation to rail vehicle dynamics. They also used an output error formulation for the estimation problem. The output error formulation is used in this thesis.

Gilan (1981) and Fries (1983) performed parameter estimation on rail vehicle models. Gilan used experimental data from on-track testing; Fries used experimental data from roller rig testing. These two works encompass much of the technology of rail vehicle parameter estimation.

Gilan used both a time domain maximum likelihood method and a frequency domain Bayesian output error method for parameter estimation. For the frequency

domain method, Gilan minimized the error between power spectral densities (PSDs) of the model and the test data. The model estimated was a lateral six degree of freedom radial truck model. The maximum likelihood method did not converge for experimental data, probably because of the inability to match the track input data to the vehicle response data. The Bayesian method did converge.

Fries used an output error parameter estimation method with Bayesian a priori parameter estimates. The model Fries estimated is a 17 degree of freedom rail vehicle called the State of the Art Car (SOAC). He implemented the estimation in the frequency domain and used frequency response functions to compare experimental and theoretical characteristics. With only slight alteration, Fries' method for estimating parameters was implemented in this thesis.

Finally, several works develop procedures for the validation of rail vehicle models. Parameter estimation is integral to this task. Fallon (1977) identified parameters to verify both linear and nonlinear mathematical models. Gilan and Hedrick (1982) evaluated lateral tangent track dynamic models by their comparison to experimental data. Gilan and Hedrick's models were developed to predict the response of passenger trucks to random alignment and crosslevel track irregularities. Gostling and Cooperrider (1983) summarized validation techniques and discussed limitations of experimental verification methods. Gostling and Cooperrider warned that "determination of the vehicle and track parameters for use in the dynamic model is one of the most critical and difficult portions of the validation process."

1.2.2 General Parameter Estimation Literature

The motions of comets and planets motivated the earliest known work in system identification and parameter estimation. Halley conducted an identification experiment in 1704 when he predicted the return of a comet from sightings of it during the previous 200 years (Norton, 1986). Gauss developed a least squares technique in the early 1800s to more accurately predict planetary orbits (Junkins, 1978). From these beginnings, parameter estimation and system identification developed into tools widely used by engineers, economists, biologists, chemists, and others concerned with modeling or control of dynamic systems.

Naturally, thousands of books and papers are available in this field. Several excellent texts, however, provide scientists and engineers enough background to use identification techniques. Junkins (1978) provided probably the most readable text on parameter estimation. This text develops theory concisely and provides a foundation to apply the estimation tool. Norton (1986) also presented a readable development of parameter estimation techniques. A text written by Ljung (1987) aims for immediate implementation of the theory. Ljung's work is unique because he developed a software toolbox operating under the program Matlab (Matlab Users Guide, 1987) to implement the theories presented. This software allows the illustration and testing of basic techniques with real and simulated data.

Other texts provide further foundation in parameter estimation. Eykhoff's (1974) classic book, though theoretically complex, includes many examples of applying identification techniques to physical problems. Texts by Sorenson (1980), Beck and Arnold (1977), and Mendel (1973) all provide general theory and algorithms for implementation.

Jain and Dobeck (1979) provided a tutorial review of identification techniques. Equation error, output error, and more advanced techniques are discussed. They considered both single-input, single-output models and multi-input, multi-output models. Trankle (1979) offered valuable advice in planning and implementing identification techniques. Trankle considered the overall technology of identification to include test planning, instrumentation specification, and choice of mathematical model, as well as numerical methods and statistical techniques of interpreting results.

Variations of Gauss' least squares method remain popular techniques today. Isenberg (1979) discussed the advantages and limitations of five classes of least squares estimators: simple, weighted, nonlinear, minimum variance, and Bayesian. Hsia (1977) also discussed these same techniques except for Bayesian estimation.

Isenberg's research is used widely in rail vehicle parameter estimation. Hasselman and Johnson (1979), Gilan (1981), and Fries (1983) all reported success using Bayesian Least Squares estimation. They also used the output error formulation and estimated parameters in the frequency domain. This implementation allowed easier interpretation of results, easier convergence of the model to the experimental data, and better noise characteristics than other methods (Fries, 1983). Thus, this thesis implements an identification method similar to the works above.

1.2.3 Identifiability Literature

Identifiability establishes that model parameters can be adequately estimated from an experiment (Norton, 1986). Certainly, identification problems accompanied parameter estimation experiments for years. However, questions regarding

identifiability arose much later than the application of parameter estimation research. As shown earlier, the first identification and estimation research was performed by Halley and Gauss. By contrast, econometricians and statisticians began working on the identifiability problem of economic models only in the early 1940s (Nguyen and Wood, 1982). Engineers, though long concerned with dynamical systems, developed identifiability applications to their field only around 1964 (Lee, 1964).

Uniform establishment of identifiability notions and definitions is a difficult task. Part of the problem is that identifiability concepts are developed simultaneously and somewhat differently in several disciplines, including physics, econometrics, biology, systems, and control (Distefano and Cobelli, 1980). Attempting to unify the contemporary concepts of identifiability, Nguyen and Wood (1982) wrote a survey paper outlining many of the existing identifiability theories. This paper also provides historical developments of the identifiability problem.

Few texts address identifiability concepts. Ljung (1987) addressed limited identifiability concepts in his textbook. Norton (1986) admitted that there are many practical difficulties with identifiability. Norton stated that no general method for global identifiability exists. Norton also wrote that "attempts at general deterministic identifiability analysis have been offered by many authors (Cobelli et al.,1979; Delforge, 1980, 1981; Walter, 1982) with varying but incomplete success (Norton, 1982b)."

Both Delforge and Norton significantly contributed to the identifiability literature (Norton, 1980, 1982a; Delforge, 1977, 1980, 1981, 1985). However, Norton and Delforge do not always agree on the results of each others research. Their differences are published in several Letters to the Editor (Delforge, 1982; Norton, 1982b).

Throughout the identifiability literature, at least six distinct theories exist (Nguyen and Wood, 1982). Of these, two identifiability methods are examined in this thesis. These methods were developed by Grewal and Glover (1976) and Reid (1979). The methods were primarily chosen for their ability to be used with standard mathematical models of rail vehicles.

Grewal and Glover's paper results from Grewal's earlier parameter estimation research (Grewal and Payne, 1976; Grewal, 1974). Though Grewal and Glover's work was originally applicable only for local identifiability, Walter et al., (1979) extended the theory to check for global identifiability. Grewal and Glover introduced the concept of output distinguishability of a structure. They then approached the identifiability problem by considering whether system outputs obtained with different parameter values can be quantitatively distinguishable from each other. These authors also showed that local output distinguishability and local least squares identifiability are completely equivalent concepts.

Reid's method is similar to the methods proposed by Norton (1986) and Delforge (1985). Reid's method uses singular value decomposition to address both identifiability and experimental design in the estimation of linear state space models. Determination of singular values/singular vectors from a computed sensitivity matrix then allows analysis of parameter identifiability and evaluation of experiment design. Reid's method is only applicable to local identifiability problems.

Identifiability is seldom mentioned in rail vehicle dynamics literature. Fries (1983) introduced a related concept of parameter estimability and defined it in terms of the parameter error obtained from a modification of the Bayesian least squares method. Unfortunately, Fries' parameter estimability test does not provide insight into the identifiability of the model structure separate from the experimental testing or data acquisition and processing.

Gilan (1981) also assessed the identifiability of various parameters from their error estimates. Gilan observed that the difficulty in estimating several of the parameters was due to the insensitivity of the cost function to the parameters. Gilan stated that the insensitivity would be independent of the estimation method he used, either the maximum likelihood method or the Bayesian least-squares method.

Finally, Hasselman and Johnson (1979) indicated difficulties in estimating the parameters of their hopper car model. They also implemented a general scheme "for assessing the statistical significance of parameter estimates." Once again, only results of the estimation program were analyzed to give an indication of the parameters' validity. This method, as with Fries' and Gilan's, did not provide a priori insight into problems with either the model structure or the measurements.

1.3 Organization of the Thesis

Selection and development of the estimation technique is presented in Chapter 2. Chapter 3 contains the development of Grewal and Glover's and Reid's identifiability techniques. Chapter 4 provides an example of the identifiability problem and implementation of the techniques developed in Chapter 3. Chapter 5 documents the experimental setup at the Transportation Test Center and details the data processing performed to obtain the experimental frequency response functions. Chapter 6 contains the development of the vertical model of the lightweight car and describes the details of the parameter estimation process. Chapter 7 summarizes and concludes the report.

Chapter 2

Estimation Technique

2.1 Introduction

The parameter estimation technique used in this thesis was an output error implementation of a weighted least-square estimation. One extra term, called the Bayesian term, was added to improve the convergence of the estimation. The method is similar to those derived by Fries (1983), Gilan (1981), and Hasselman and Johnson (1979). This method minimizes a quadratic performance index containing the differences between the measured and expected outputs. The estimation is performed in the frequency domain because it allows easier interpretation of the results and easier matching of the experiment and the model.

This chapter contains the selection and derivation of the estimation technique used to determine the parameters of a 3 degree-of-freedom symmetrical rail vehicle model. The choice of the estimation technique is straightforward: the output error

method was successful in previous rail vehicle parameter estimation research and its implementation is well documented. However, completeness still merits the examination of several of the most popular estimation techniques.

2.2 *Selection of Estimation Technique*

2.2.1 Summary of Popular Estimation Techniques

Three major estimation techniques are presented in the literature:

- equation error minimization methods
- output error minimization methods
- simultaneous state and parameter estimation methods

Several authors document the development and application of these methods (Trankle, 1979; Jain and Dobeck, 1979; Fries, 1983). Each method has distinct advantages and disadvantages which are briefly discussed below. This development follows closely the work of Trankle and Fries.

2.2.1.1 Equation Error Methods

A continuous dynamic system is represented as:

$$\dot{\underline{x}} = \frac{d\underline{x}}{dt} = \underline{F}(\underline{x}, \underline{\theta}, \underline{u}, t) + \underline{\eta}(t) \qquad [2.2.1]$$

where

$\underline{\theta}$ = a set of p unknown parameters

$\underline{\eta}$ = a time-varying unobservable disturbance vector

$\underline{x}$ = system states

t = time

The equation error method finds a parameter vector, $\underline{\theta}$ to minimize a quadratic performance index composed of the difference between the measured and expected state derivatives

$$J_e = (\hat{\dot{\underline{x}}} - \underline{F})' (\hat{\dot{\underline{x}}} - \underline{F}) \qquad [2.2.2]$$

or

$$J_e = \sum_{i=1}^{M} \left[\frac{d\hat{\underline{x}}(t_i)}{dt} - F(\hat{\underline{x}}(t_i), \hat{\underline{u}}(t_i), t_i, \underline{\theta}) \right]^2 \qquad [2.2.3]$$

where M is the number of measurements, the $\wedge$ denotes a measured value, and F is a computed value. The parameter vector $\underline{\theta}$ is determined by solving

$$\frac{\partial J_e}{\partial \underline{\theta}} = \underline{0}. \qquad [2.2.4]$$

An advantage of the equation error method is its effectiveness in the presence of process noise (noise due to inadequate modeling). A further advantage of the equation error method is that many nonlinear dynamic system functions $\underline{F}(\underline{x}, \underline{u}, \underline{\theta}, t)$ are linear in the parameters $\underline{\theta}$. This property may be represented by

$$\underline{E}(\underline{x}, \underline{u}, \underline{\theta}, t) = \sum_{j=1}^{p} \theta_j \underline{F}_j(\underline{x}, \underline{u}, t) + \underline{F}_{p+1}(\underline{x}, \underline{u}, t). \qquad [2.2.5]$$

The functions F_j , $j = 1,2, ..., p + 1$ are independent of the p unknown parameters θ_j . The parameter values which minimize the performance index can be found explicitly using linear algebraic operations. This property makes the equation error method computationally simple.

The disadvantage of the equation error method is that it requires the a priori determination of all the system states, $\underline{x}$, inputs $\underline{u}$, and state derivatives $\dot{\underline{x}}$. For the experiments performed at the Transportation Test Center, only the states of the system were measured; further computations would be needed to determine the state derivatives.

2.2.1.2 Output Error Methods

For the output error method, the dynamic system is represented by

$$\dot{\underline{x}} = \underline{F}(\underline{x}, \underline{\theta}, \underline{u}, t) \qquad [2.2.6]$$

and

$$\underline{y} = \underline{H}(\underline{x}, \underline{\theta}, \underline{u}, t) + \underline{\eta}(t). \qquad [2.2.7]$$

The solution is obtained by minimizing a quadratic performance index containing the differences between the measured and expected outputs. The performance index is represented as

$$J_o = (\hat{\underline{y}} - \tilde{y}(t_i, \underline{\theta}))' (\hat{\underline{y}} - \tilde{y}(t_i, \underline{\theta})) \qquad [2.2.8]$$

or

$$J_o = \sum_{i=1}^{M} [\hat{y}_i - \tilde{y}(t_i, \underline{\theta})]^2. \qquad [2.2.9]$$

Here, $\hat{\underline{y}}$ is the observed system output and $\tilde{y}(t_i, \underline{\theta})$ is the predicted system output obtained by solving the system state equations and output equations from the measured system inputs $\hat{\underline{u}}(t_i)$ and the a priori parameter values $\underline{\theta}$.

There are several advantages of the output error method. First, the output error method significantly relaxes the measurement requirements compared to the equation error method because the state derivatives are not required. Second, the output error method is effective in the presence of measurement errors. A disadvantage of the output error method is its sensitivity to process noise.

2.2.1.3 Simultaneous State and Parameter Estimation Methods

Methods that estimate both the system states and parameters are not biased by either process or measurement noise. The performance index here is similar to the output error performance index. The performance index for the combined state and parameter estimation method is

$$J_s = (\hat{\underline{y}} - \tilde{y}(t_i, \underline{\theta}, \hat{\underline{y}}))' (\hat{\underline{y}} - \tilde{y}(t_i, \underline{\theta}, \hat{\underline{y}})) \qquad [2.2.10]$$

or

$$J_s = \sum_{i=1}^{M} [\hat{y}_i - \tilde{y}(t_i, \underline{\theta}, \hat{\underline{y}})]^2. \qquad [2.2.11]$$

The major difference between the output error method and the combined state and parameter error method is the estimated outputs, $\tilde{\underline{y}}$, are now also direct functions of the observed outputs, $\hat{\underline{y}}$. Naturally, the benefits of this estimation method are offset by several disadvantages. The simultaneous estimation is more complex, more computing intensive, and less stable than the other two methods.

Gilan (1981) implemented a combined state and parameter estimation method. The method, called the maximum likelihood method, is well developed and documented. Gilan successfully implemented the maximum likelihood estimation using simulated data. However, for real on-track data, the estimation process would not converge. A probable cause for this lack of convergence is the difficulty Gilan met trying to match the time references of the track input data and the vehicle response data.

2.2.2 Selection of Output Error Estimation Technique

The three basic parameter estimation methods are summarized in Table 1. Two criteria will be weighed to choose the estimation process:

1. Conclusions from literature
2. Noise characteristics

Table 1. Parameter Estimation Methods

Method	*Advantages*	*Disadvantages*	*Performance Index*
Equation Error	Effective in presence of process noise Computational simplicity	Sensitive to measurement errors	$\sum(\dot{\hat{x}} - \tilde{F}(\hat{\underline{x}}, \hat{\underline{u}}, \underline{\theta}, t))^2$
Output Error	Effective in presence of measurement errors	Sensitive to process noise	$\sum(\hat{y} - \tilde{y}(\underline{\theta}, t))^2$
Combined State and Parameter Estimation	Effective in presence of both measurement and process noise	Computational complexity	$\sum(\hat{y} - \tilde{y}(\underline{\theta}, \hat{\underline{y}}, t))^2$

2.2.2.1 Conclusions from Literature

A significant justification for implementing the output error method is the strong supporting body of literature. Fries (1983), Gilan (1981), and Hasselman and Johnson (1979) all concluded that the output error method was an efficient estimator for rail vehicle dynamics models. Trankle (1979) cited work by Broersen (1973) that

implemented the output error method for rail vehicle dynamics. Gostling and Cooperrider expressed a common opinion of researchers by stating: "Methods which minimize output errors (i.e. the difference between measured and predicted outputs) have emerged as the most workable approaches to railway vehicle parameter identification."

Gilan concluded, even though he performed a maximum likelihood estimation also, that the output error method was well suited for the estimation. The maximum likelihood method is very sophisticated and is often difficult to apply. Gilan recommended the output error method and stated that it was "simpler to implement, less restrictive, and less time consuming."

Of course, each parameter estimation experiment must be evaluated on its own merits, and not solely on conclusions from previous work. Therefore, the specific experiment in this thesis must be examined to determine the optimum estimation method.

2.2.2.2 Noise Characteristics

An understanding of the experiment is necessary to choose an estimation method. The quantity and quality of the measurements affects which estimation method will be most successful.

The lightweight rail vehicle was tested on a large hydraulically actuated test unit called the Vibration Test Unit (VTU). The VTU is capable of providing accurate input waveforms to the wheels of the car through a rail beam mounted on the actuator. The input waveform contains very little noise, and most of the measured channels are fairly clean also. The clean input would benefit all of the estimation

methods but would substantially benefit the output error method because the predicted output term is a function of only the measured inputs and model structure.

The vehicle experiments did not contain any measurements of the state derivatives. In some instances, the states and state derivatives of the model do not even have any physical significance. Because of this, the equation error method becomes more difficult to implement. When information on the state derivatives is not available, it is sometimes possible to use numerical techniques to obtain the missing state derivatives. This difficulty decreases the utility of the equation error technique.

Thus, taking into consideration much of the previous rail vehicle parameter estimation research and examining the properties of the specific experiment performed, the output error method emerged as the best technique. In the next sections, the output error estimation algorithm will be presented for the vehicle model in the frequency domain using frequency response functions.

2.3 Development of the Output Error Estimation Technique

The output error estimation method can be implemented in either the frequency domain or the time domain. In the previous section, it was introduced as a time domain method. The domain choice for the estimation is not necessarily direct.

Ljung and Glover (1981) considered frequency domain versus time domain methods in system identification. They outlined several techniques adapting frequency and time domains and explored the connections and distinctions between these methods. A main conclusion of their work was that "the two approaches (time domain and frequency domain) are complementary rather than rivalling."

The purpose of the estimation process must also be examined to determine which domain is more appropriate. The purpose of the estimation here is to provide engineers with a valid linear mathematical model for the rail vehicle. This model will then be used to evaluate the safety aspects of the new vehicle designs. Ljung and Glover stated that "the object of the identification may be to simply gain general insight into the system, determining for example resonances in the response. Here frequency domain techniques are probably most appropriate." This statement applies directly to the purpose of this thesis. Thus, the estimation was performed in the frequency domain.

2.3.1 Frequency Domain Implementation of Output Error Method

This section shows the formulation of the output error estimation problem in the frequency domain. This development follows the work by Fries (1983).

Assume a linear, time-invariant system of the form

$$\dot{\underline{x}} = \mathbf{A}\underline{x} + \mathbf{B}\underline{u} \qquad [2.3.1]$$

$$\underline{y} = \mathbf{C}\underline{x} + \mathbf{D}\underline{u} + \underline{\eta}. \qquad [2.3.2]$$

Transforming to the frequency domain, these equations become

$$j\omega \underline{X} = \mathbf{A}\underline{X} + \mathbf{B}\underline{U} \qquad [2.3.3]$$

$$\underline{Y} = \mathbf{C}\underline{X} + \mathbf{D}\underline{U} + \underline{\eta}^{\ddagger} \qquad [2.3.4]$$

where the capital letters and the ‡ denote the Fourier transform of the function, ω is the frequency, and $j = \sqrt{-1}$. From equation [2.3.3],

$$(j\omega \mathbf{I} - \mathbf{A})\,\underline{X} = \mathbf{B}\underline{U} \qquad [2.3.5]$$

and

$$\underline{X} = (j\omega \mathbf{I} - \mathbf{A})^{-1}\,\mathbf{B}\,\underline{U}. \qquad [2.3.6]$$

Substituting into equation [2.3.4] gives

$$\underline{Y} = [\mathbf{C}\,(j\omega \mathbf{I} - \mathbf{A})^{-1}\,\mathbf{B} + \mathbf{D}]\,\underline{U} + \underline{\eta}^{\ddagger}. \qquad [2.3.7]$$

This equation relates the Fourier transform of the output, $\underline{Y}$, to the Fourier transform of the input $\underline{U}$ and the Fourier transform of the disturbance $\underline{\eta}^{\ddagger}$.

To form the output error estimation method, the performance index is written as

$$J = (\hat{\underline{Y}} - \underline{Y})'\,\mathbf{W}\,(\hat{\underline{Y}} - \underline{Y}) \qquad [2.3.8]$$

where $\hat{\underline{Y}}$ is the Fourier transform of the measurement output and **W** is the observation weight matrix discussed by Isenberg (1979).

The observation weight matrix, **W**, is an extension of the least-squares estimation. The observation weight matrix accounts for both measurement of data with different precision and incorporation of data with different dimensions. This is

accomplished by incorporating statistical properties of the data into the estimation algorithm.

The performance index can also be formulated with measured PSDs or frequency response functions (FRFs). In this work, as in Fries (1983), frequency response functions provided the basis for comparison. Therefore, the performance index was

$$J = (\hat{\underline{H}} - \underline{H})' \mathbf{W} (\hat{\underline{H}} - \underline{H}) \qquad [2.3.9]$$

where $\hat{\underline{H}}$ is the set of frequency response functions computed from the measured outputs and $\underline{H}$ is the corresponding set of model frequency response functions. The weighting matrix, **W**, is formulated specifically for frequency response function error and is shown in a later section. Note that the theoretical FRF matrix, **H**, is obtained from

$$\mathbf{H} = [\mathbf{C} (j\omega \mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D}]. \qquad [2.3.10]$$

A disadvantage of the formulation is the parameters are generally not linearly related to the FRFs. Even if **A**, **B**, **C**, and **D** from equation [2.3.10] are linear in the parameters, **H** is generally nonlinear in the parameters. This requires the solution of a nonlinear set of algebraic equations to determine the parameter estimates.

2.3.2 Minimum Variance Output Error Estimation Algorithm

This section contains a derivation of the minimum variance output error estimation algorithm. The derivation includes the addition of a priori parameter estimates, and is usually called Bayesian estimation. Isenberg (1979) provided a

review of the five basic classes of least-square estimators. The most advanced technique presented by Isenberg was the Bayesian minimum variance least-square estimator.

The minimum variance estimator is a general tool for parameter estimation. However, for many measurements (observations) the weighting matrix, being of size n x n , becomes prohibitively large for implementation on a computer. This problem can be resolved by dividing the experimental data into batches and processing them sequentially. However, to accomplish this, there needs to be some method of communicating the results from any subset of the estimation to the remainder of the estimation process. The Bayesian method accomplishes this purpose.

The Bayesian estimation allows independent processing of the observation sets. Fries (1983) illustrated the interaction of the parameter sets. For example, there might be three sets of observations to be processed. The first set could be processed independently of the second and third by using the minimum variance algorithm without prior estimates. This algorithm would provide the parameter estimates, $\underline{\theta}$, and a measure of the quality of the estimates. The quality measure is the parameter error covariance matrix, $\mathbf{S}_{rr}$. With this information the second and third sets of observations could be processed using the estimates and covariance error matrix from the first stage of the processing. To incorporate this information into the estimation process, the performance index for the second and each following stage of the estimation process is written as

$$J = (\hat{\underline{y}} - \underline{y})' \mathbf{S}_{\varepsilon\varepsilon}^{-1} (\hat{\underline{y}} - \underline{y}) + (\underline{\theta}_o - \underline{\theta})' \mathbf{S}_{rr}^{-1} (\underline{\theta}_o - \underline{\theta}) \qquad [2.3.11]$$

where

$\hat{\underline{y}}$ = the measurement vector

$\underline{y}$ = the computed system output vector

S_{zz} = the observation error covariance matrix

$\underline{\theta}_o$ = the prior parameter error vector

$\underline{\theta}$ = the current parameter estimate vector

and

S_{rr} = the parameter error covariance matrix.

2.3.2.1 Derivation

For this derivation, the system output equation [2.3.7] is written as

$$\underline{y} = \underline{H}(\underline{x}, \underline{\theta}) \qquad [2.3.12]$$

where

$\underline{x}$ = the state vector

$\underline{\theta}$ = the parameter vector

and

$\underline{H}$ = the output equation set.

Because $\underline{H}$ is usually nonlinear in the parameter vector $\underline{\theta}$, an iterative solution is expected, and the performance index will be written in terms of its revised value at the end of an iteration. Therefore, let

$$\underline{\theta}^{\dagger} = \underline{\theta} + \Delta\underline{\theta} \qquad [2.3.13]$$

where

$\underline{\theta}^{\dagger}$ = the revised estimate of $\underline{\theta}$

and

$\Delta\underline{\theta}$ = the change in $\underline{\theta}$ from the current to the next step

The revised system output is then approximated by the first two terms of the Taylor series expansion of the system equation.

$$\underline{y}^{\dagger} = \underline{y} + \frac{\partial \underline{y}}{\partial \underline{\theta}} \Delta\underline{\theta} \qquad [2.3.14]$$

or

$$\underline{y}^{\dagger} = \underline{y} + \mathbf{T}\,\Delta\underline{\theta} \qquad [2.3.15]$$

where

$$T_{ij} = \frac{\partial y_i}{\partial \theta_j} \qquad [2.3.16]$$

or, in matrix form,

$$\mathbf{T} = \frac{\partial \underline{y}}{\partial \underline{\theta}}. \qquad [2.3.17]$$

The revised performance index is

$$J^{\dagger} = (\hat{\underline{y}} - \underline{y}^{\dagger})' \, \mathbf{S}_{\varepsilon\varepsilon}^{-1} \, (\hat{\underline{y}} - \underline{y}^{\dagger}) + (\underline{\theta}_o - \underline{\theta}^{\dagger})' \, \mathbf{S}_{rr}^{-1} \, (\underline{\theta}_o - \underline{\theta}^{\dagger}). \qquad [2.3.18]$$

Substituting equation [2.3.13] and [2.3.15] into [2.3.18], gives the performance index in terms of the linearized system equations.

$$J^{\dagger} = (\hat{\underline{y}} - \underline{y} - \mathbf{T}\Delta\underline{\theta})' \, \mathbf{S}_{\varepsilon\varepsilon}^{-1} \, (\hat{\underline{y}} - \underline{y} - \mathbf{T}\Delta\underline{\theta}) + (\underline{\theta}_o - \underline{\theta} - \Delta\underline{\theta})' \, \mathbf{S}_{rr}^{-1} \, (\underline{\theta}_o - \underline{\theta} - \Delta\underline{\theta}). \qquad [2.3.19]$$

A value for the parameter correction vector, $\Delta\underline{\theta}$, that minimizes $J^{\dagger}$ is obtained by

$$\frac{\partial J^{\dagger}}{\partial \Delta\underline{\theta}} = \underline{0}. \qquad [2.3.20]$$

This partial differentiation leads to

$$-\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}(\hat{\underline{y}}-\underline{y}-\mathbf{T}\Delta\underline{\theta})-\mathbf{S}_{rr}^{-1}(\underline{\theta}_o-\underline{\theta}-\Delta\underline{\theta})=0. \qquad [2.3.21]$$

Therefore,

$$(\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T}+\mathbf{S}_{rr}^{-1})\Delta\underline{\theta} = \mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}(\hat{\underline{y}}-\underline{y})+\mathbf{S}_{rr}^{-1}(\underline{\theta}_o-\underline{\theta}) \qquad [2.3.22]$$

or

$$\Delta\underline{\theta} = [(\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T}+\mathbf{S}_{rr}^{-1})]^{-1}[\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}(\hat{\underline{y}}-\underline{y})+\mathbf{S}_{rr}^{-1}(\underline{\theta}_o-\underline{\theta})]. \qquad [2.3.23]$$

Writing this equation in terms of the revised parameter estimates gives

$$\underline{\theta}^{\dagger} = \underline{\theta}+[(\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T}+\mathbf{S}_{rr}^{-1})]^{-1}[\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}(\hat{\underline{y}}-\underline{y})+\mathbf{S}_{rr}^{-1}(\underline{\theta}_o-\underline{\theta})]. \qquad [2.3.24]$$

Equation [2.3.24] is the general minimum variance parameter estimation update equation. This equation is applied repeatedly until $\underline{\theta}^{\dagger}$ converges within an acceptable tolerance.

This estimation method also provides a measure of the quality of the parameter estimates. To determine the revised observation error covariance matrix, $\mathbf{S}_{rr}$, results from the linear system case are used. When $\underline{\mathbf{H}}$ is a linear function of the parameters, $\underline{\theta}$, equation [2.3.24] will converge to the global minimum of J in one step. For the linear case, equation [2.3.24] can therefore be written as

$$\underline{\theta}^{\dagger} = [(\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T}+\mathbf{S}_{rr}^{-1})]^{-1}[\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}(\hat{\underline{y}}-\underline{y})+\mathbf{S}_{rr}^{-1}(\underline{\theta}_o-\underline{\theta})] \qquad [2.3.25]$$

so $\underline{\theta}^{\dagger}$ is linearly related to the observation error $(\hat{\underline{y}}-\underline{y})$ and the parameter error $(\underline{\theta}_o-\underline{\theta})$.

To determine the revised observation error covariance matrix, the random error in the observation vector and the parameter vector must be considered. Therefore, if

$Z_{rr}^{\dagger}$ = the random error in $\underline{\theta}^{\dagger}$

and

$Z_{\varepsilon\varepsilon}$ = the random error in the measurement

then

$$\mathbf{S}_{rr}^{\dagger} = E(Z_{rr}^{\dagger} Z_{rr}^{\dagger\prime}) \qquad [2.3.26]$$

and

$$\mathbf{S}_{\varepsilon\varepsilon} = E(Z_{\varepsilon\varepsilon} Z_{\varepsilon\varepsilon}^{\prime}) \qquad [2.3.27]$$

where

E = the expected value operator.

Therefore, $Z_{rr}^{\dagger}$ is obtained directly from equation [2.3.25] :

$$Z_{rr}^{\dagger} = [\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1}]^{-1} [\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}Z_{\varepsilon\varepsilon} + \mathbf{S}_{rr}^{-1}Z_{rr}]. \qquad [2.3.28]$$

Substituting equation [2.3.28] into equation [2.3.26] gives

$$\begin{aligned}\mathbf{S}_{rr}^{\dagger} = {} & [\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1}]^{-1}\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}E(Z_{\varepsilon\varepsilon}Z_{\varepsilon\varepsilon}')[(\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1})^{-1}\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}]' \\ & + [\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1}]^{-1}\mathbf{S}_{rr}^{-1}E(Z_{rr}Z_{rr}')[(\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1})^{-1}\mathbf{S}_{rr}^{-1}]'\end{aligned} \qquad [2.3.29]$$

where it is assumed that the covariance between the parameters and outputs is zero, or

$$E(Z_{rr}Z_{\varepsilon\varepsilon}') = E(Z_{\varepsilon\varepsilon}Z_{rr}') = 0 \qquad [2.3.30]$$

Now,

$$\mathbf{S}_{\varepsilon\varepsilon}^{-1}E(\mathbf{Z}_{\varepsilon\varepsilon}\mathbf{Z}_{\varepsilon\varepsilon}') = \mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{S}_{\varepsilon\varepsilon} = \mathbf{I} \qquad [2.3.31]$$

and

$$\mathbf{S}_{rr}^{-1}E(\mathbf{Z}_{rr}\mathbf{Z}_{rr}') = \mathbf{S}_{rr}^{-1}\mathbf{S}_{rr} = \mathbf{I}. \qquad [2.3.32]$$

Therefore, equation [2.3.29] can be rewritten as

$$\begin{aligned}\mathbf{S}_{rr}^{\dagger} = {} & (\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1})^{-1}[(\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1})^{-1}\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T}]' \\ & + (\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1})^{-1}[(\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1})^{-1}\mathbf{S}_{rr}^{-1}]'.\end{aligned} \qquad [2.3.33]$$

Collecting terms gives

$$\mathbf{S}_{rr}^{\dagger} = (\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1})^{-1}[(\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1})^{-1}(\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1})]'. \qquad [2.3.34]$$

This equation finally reduces to

$$\mathbf{S}_{rr}^{\dagger} = (\mathbf{T}'\mathbf{S}_{\varepsilon\varepsilon}^{-1}\mathbf{T} + \mathbf{S}_{rr}^{-1})^{-1}. \qquad [2.3.35]$$

Therefore, equation [2.3.35] provides a method to update the parameter error covariance matrix at each stage of the estimation process. This equation is an approximation for the nonlinear case, since only the first two terms of the Taylor series approximation were used. The validity of equation [2.3.35] also requires that the original estimate of the measurement error covariance matrix is correct; otherwise, equation [2.3.27] is not satisfied.

2.3.2.2 *Implementation*

Fries (1983) recommended the following algorithm when estimating large linear dynamic models.

1. Estimate $\mathbf{S}_{rr}$.
2. Estimate $\underline{\theta}_o$.
3. For the first set of observations to be processed, let $\underline{\theta} = \underline{\theta}_o$ and $\mathbf{S}_{rr}^{-1} = [\mathbf{0}]$ in equation [2.3.24] and compute $\underline{\theta}^\dagger$.
4. Let $\underline{\theta} = \underline{\theta}^\dagger$ and compute a new $\underline{\theta}^\dagger$ from equation [2.3.24].
5. Continue the iteration in step 4 until the process converges.
6. Compute $\mathbf{S}_{rr}^\dagger$ using equation [2.3.35], keeping $\mathbf{S}_{rr}^{-1} = [\mathbf{0}]$.
7. Process the next set of observations using $\underline{\theta}_o = \underline{\theta}^\dagger$ and $\mathbf{S}_{rr} = \mathbf{S}_{rr}^\dagger$ from the first set of observations processed.
8. Continue step 7 until all the observation sets are processed.

Isenberg (1979) showed that for a linear system the above sequential processing procedure gives the same results that would be obtained by setting $\mathbf{S}_{rr}^{-1} = [\mathbf{0}]$ and processing all of the observations simultaneously. For a nonlinear system the results would be similar, but not exact for the two methods.

Finally, the above algorithm may be altered slightly to incorporate more knowledge of the physical parameters and to improve the convergence of the algorithm. For example, if one of the physical parameters such as mass is known with more confidence than other parameters, the corresponding value for its parameter error covariance in the term $\mathbf{S}_{rr}$ may be chosen as a much smaller value than for the other parameters. This alteration would yield smaller changes in the mass compared to changes in other parameters such as stiffness or damping.

However, if this alteration is not used cautiously, inappropriate restrictions on parameter values are imposed and the estimation will converge to incorrect values.

Chapter 3

Identifiability

3.1 Introduction

Identifiability establishes that model parameters can be adequately estimated from the experiment (Norton, 1986). Identifiability is generally independent of the parameter estimation method used. However, both model structure and parameter values affect identifiability.

Identifiability problems emerge in several disciplines. Biologists researching cell dynamics encounter problems with identifiability. Economists estimating parameters for econometric modeling meet difficulty also. However, within mechanical engineering and especially in rail vehicle dynamics research there is little emphasis in the literature on the identifiability of the dynamic models.

A possible cause for the lack of investigation is that identifiability problems usually hide within the parameter estimation task of mathematical modeling.

Therefore, the problem is seen to emanate from the estimation algorithm and not the model structure.

Suspected examples of identifiability problems are seen in literature on estimation of rail vehicle dynamics. One example of this problem comes from Hasselman and Johnson's (1979) paper. The authors stated, "It is entirely possible that the parameters of a model may be adjusted so as to achieve a closer match between predicted response and measured response with no resultant improvement in the model." They also concluded, "If one lesson stands out from the authors' experience to date, it is that a model may be adjusted to match certain data without actually improving the model, e.g. the wrong parameters may be changed."

3.2 *Identifiability Techniques*

Throughout the identifiability literature, at least six distinct theories exist. Of these, two identifiability methods are examined in detail. These methods were developed by Grewal and Glover (1976) and Reid (1979). The methods were primarily chosen for their ability to be used with standard mathematical models of rail vehicles. Both methods use the state-variable description of the system. The state-variable technique enables the physical significance of the parameters to be retained. Both of the methods are used to check for local identifiability only. Walter et al. (1979), however, extended Grewal and Glover's method for global identifiability.

3.2.1 Grewal and Glover's Method

Grewal and Glover's (1976) method for identifiability of model structures considers if outputs obtained with identical system structure but different parameter values are unique. They developed the idea of output distinguishability and defined parameter identifiability using output distinguishability.

3.2.1.1 Formulation

A parameterized system may be described by

$$\dot{\underline{x}} = \mathbf{A}(\underline{\theta})\underline{x} + \mathbf{B}(\underline{\theta})\underline{u} \qquad [3.2.1]$$

with output equations

$$\underline{y}(t) = \mathbf{C}(\underline{\theta})\,\underline{x}(t) + \mathbf{D}(\underline{\theta})\underline{u} \qquad [3.2.2]$$

and initial conditions

$$\underline{x}(0) = \underline{0}. \qquad [3.2.3]$$

To develop Grewal and Glover's method, first let

Ω = parameter space for all parameters of the system

$\underline{\theta}$ = parameter vector in Ω

$\underline{\alpha}$ = parameter vector in Ω

Grewal and Glover showed a parameter set Ω is identifiable at $\underline{\theta}$ if the pair $(\underline{\alpha}, \underline{\theta})$ is distinguishable for all $\underline{\alpha} \in \Omega$, $\underline{\alpha} \neq \underline{\theta}$. Formulating the identifiability condition, Grewal

and Glover first formed the general solution of the state variable equations. If ($\underline{\theta}$, $\underline{\alpha}$) are indistinguishable, then

$$\mathbf{C}(\underline{\theta})\int_0^t e^{\mathbf{A}(\underline{\theta})(t-\tau)}\,\mathbf{B}(\underline{\theta})u(\tau)d\tau + \mathbf{D}(\underline{\theta})u(t) \equiv \mathbf{C}(\underline{\alpha})\int_0^t e^{\mathbf{A}(\underline{\alpha})(t-\tau)}\,\mathbf{B}(\underline{\alpha})u(\tau)d\tau + \mathbf{D}(\underline{\alpha})u(t) \quad [3.2.4]$$

This relationship requires both

$$\mathbf{D}(\underline{\theta}) = \mathbf{D}(\underline{\alpha}) \quad [3.2.5]$$

and

$$\int_0^t [\mathbf{C}(\underline{\theta})e^{\mathbf{A}(\underline{\theta})(t-\tau)}\,\mathbf{B}(\underline{\theta}) - \mathbf{C}(\underline{\alpha})e^{\mathbf{A}(\underline{\alpha})(t-\tau)}\,\mathbf{B}(\underline{\alpha})]u(\tau)d\tau \equiv 0. \quad [3.2.6]$$

Using a lemma from calculus of variations (Gelphan and Fomin, 1963), the equation above becomes

$$\mathbf{C}(\underline{\theta})e^{\mathbf{A}(\underline{\theta})(t-\tau)}\,\mathbf{B}(\underline{\theta}) - \mathbf{C}(\underline{\alpha})e^{\mathbf{A}(\underline{\alpha})(t-\tau)}\,\mathbf{B}(\underline{\alpha}) = 0 \quad [3.2.7]$$

for all $0 \leq \tau \leq t$. By repeated differentiation of the time invariant response

$$\begin{aligned}
\mathbf{C}(\underline{\theta})\,\mathbf{B}(\underline{\theta}) &= \mathbf{C}(\underline{\alpha})\,\mathbf{B}(\underline{\alpha}) \\
\mathbf{C}(\underline{\theta})\,\mathbf{A}(\underline{\theta})\,\mathbf{B}(\underline{\theta}) &= \mathbf{C}(\underline{\alpha})\,\mathbf{A}(\underline{\alpha})\,\mathbf{B}(\underline{\alpha}) \\
&\vdots \qquad\qquad \vdots \\
\mathbf{C}(\underline{\theta})\,\mathbf{A}(\underline{\theta})^{\ell}\,\mathbf{B}(\underline{\theta}) &= \mathbf{C}(\underline{\alpha})\,\mathbf{A}(\underline{\alpha})^{\ell}\,\mathbf{B}(\underline{\alpha})
\end{aligned} \quad [3.2.8]$$

where

$\ell = 0,1,2,...$, number of parameters to identify

Therefore, indistinguishability implies that at different values of the parameter vector, $\underline{\theta}$, the parameters of the system $\mathbf{CA^{\ell}B}$ are identical. The resulting theorem is presented as described by Grewal and Glover.

Theorem: A parameter set Ω is identifiable at $\underline{\theta}$ if and only if

$$\underline{\alpha} \in \Omega, \qquad [3.2.9]$$

$$\mathbf{D}(\underline{\theta}) = \mathbf{D}(\underline{\alpha}), \qquad [3.2.10]$$

and

$$\mathbf{C}(\underline{\theta})\,\mathbf{A}(\underline{\theta})^{\ell}\,\mathbf{B}(\underline{\theta}) = \mathbf{C}(\underline{\alpha})\,\mathbf{A}(\underline{\alpha})^{\ell}\,\mathbf{B}(\underline{\alpha}), \quad \ell = 0,1,2,... \qquad [3.2.11]$$

together imply

$$\underline{\alpha} = \underline{\theta}. \qquad [3.2.12]$$

3.2.1.2 Identifiability Condition

One approach to implementing the identifiability theorem is to consider the Markov parameter matrix

$$\mathbf{G}(\underline{\theta}) = \begin{bmatrix} \mathbf{D}(\underline{\theta}) \\ \mathbf{C}(\underline{\theta})\,\mathbf{B}(\underline{\theta}) \\ \mathbf{C}(\underline{\theta})\,\mathbf{A}(\underline{\theta})\,\mathbf{B}(\underline{\theta}) \\ \mathbf{C}(\underline{\theta})\,\mathbf{A}^2(\underline{\theta})\,\mathbf{B}(\underline{\theta}) \\ \vdots \\ \mathbf{C}(\underline{\theta})\,\mathbf{A}^{\ell}(\underline{\theta})\,\mathbf{B}(\underline{\theta}) \end{bmatrix} \qquad [3.2.13]$$

where

ℓ = 1, 2, 3, ..., number of parameters to identify

and check if it is one to one on Ω at $\hat{\underline{\theta}}$, the local values of $\underline{\theta}$. The mapping from the parameter space to the Markov parameter space is locally one to one if the rank of the Jacobian of $\mathbf{G}(\underline{\theta})$ is equal to the number of the parameters to identify. The Jacobian of the Markov parameter matrix is

$$\frac{\partial \mathbf{G}(\underline{\theta})}{\partial \underline{\theta}_i} = \begin{bmatrix} \frac{\partial}{\partial \theta_i}[\mathbf{D}(\underline{\theta})] \\ \frac{\partial}{\partial \theta_i}[\mathbf{C}(\underline{\theta})\,\mathbf{B}(\underline{\theta})] \\ \frac{\partial}{\partial \theta_i}[\mathbf{C}(\underline{\theta})\,\mathbf{A}(\underline{\theta})\,\mathbf{B}(\underline{\theta})] \\ \frac{\partial}{\partial \theta_i}[\mathbf{C}(\underline{\theta})\,\mathbf{A}^2(\underline{\theta})\,\mathbf{B}(\underline{\theta})] \\ \vdots \\ \frac{\partial}{\partial \theta_i}[\mathbf{C}(\underline{\theta})\,\mathbf{A}^{\ell}(\underline{\theta})\,\mathbf{B}(\underline{\theta})] \end{bmatrix} \qquad [3.2.14]$$

Therefore, for the parameters of a dynamic system to be identifiable, the rank of equation [3.2.14] must be equal to or greater than the number of estimated parameters.

3.2.2 Reid's Method

The method for determining the identifiability of a model structure developed by Reid (1979) differs from Grewal and Glover's method (1976) primarily by the use of a matrix of eigenvalue and eigenvector sensitivities. Reid's method also uses singular value decomposition to address both identifiability and experimental design in the parameter identification of linear state space models.

Reid's method investigates both the identifiability of a model structure and the specific experimental design quality. This work considered only the identifiability issues. Combining identifiability and experimental design issues, however, is essential in developing identification experiments. For example, increasing the amount of data taken or increasing the control energy to a system does not improve the parameter estimation results if the structural requirements on the model are not met.

3.2.2.1 Formulation

Reid's method also uses the state variable description of a dynamic system. First, assume a system of the form

$$\dot{\underline{x}}(\underline{\theta}) = \mathbf{A}(\underline{\theta})\underline{x} + \mathbf{B}(\underline{\theta})\underline{u} \qquad [3.2.15]$$

with output equations

$$\underline{y}(t_k) = \mathbf{C}(\underline{\theta})\,\underline{x}(t_k) \qquad [3.2.16]$$

and initial conditions

$$\underline{x}(0) = \underline{x}_o(\underline{\theta}). \qquad [3.2.17]$$

The dimensions of the matrices are:

n = the state dimension

q = the output dimension

p = the number of unknown parameters, θ_i

k = the number of discrete output sample times, t_k

The output dimension is one.

Assuming **A** has n distinct eigenvalues, the solution at any time t_k is

$$\underline{y}(t_k) = \mathbf{C}e^{\mathbf{A}t_k}\underline{x}_o + \mathbf{C}\int_0^{t_k} e^{\mathbf{A}(t_k-\tau)}\mathbf{B}\underline{u}(\tau)d\tau \qquad [3.2.18]$$

This equation was rewritten by Reid as

$$\underline{y}(t_k) = \sum_{j=1}^{n} \mathbf{C}\,[\underline{v}_j\,\underline{w}_j{}']\,\underline{x}_o\,e^{\lambda_j t_k} + \sum_{j=1}^{n} \mathbf{C}\,[\underline{v}_j\,\underline{w}_j{}']\,\mathbf{B}\,e^{\lambda_j t_k}\int_0^{t_k} e^{-\lambda_j\tau}\,\underline{u}(\tau)\,d\tau \qquad [3.2.19]$$

where $\underline{v}_j$ and $\underline{w}_j$ are the right and left eigenvectors corresponding to the eigenvalues λ_j , j = 1,2,...,n. The eigenvector matrices, **v** and **w**, are related by

$$\mathbf{w} = [\mathbf{v}']^{-1} \qquad [3.2.20]$$

Reid then defined the parameter sensitivity matrix, **S** , by forming the partial derivative of [3.2.19] at each time component t_k , k = 1, 2 , ...,k , with respect to each parameter component θ_i , i = 1, 2, ..., p . With this operation, the sensitivity matrix **S** decomposes into its structural and time dependent parts as

$$\mathbf{S} = [\mathbf{E}, \mathbf{F}]\mathbf{G} \qquad [3.2.21]$$

This procedure reduces the partial derivative of equation [3.2.19] into less complex matrices, grouping the time dependent and structural portions. Each portion of **S** can be expressed in more detail. Beginning with the **G** matrix,

$$\mathbf{G} \equiv \begin{bmatrix} \mathbf{G}_{zi} \\ \mathbf{G}_{zs} \end{bmatrix}_{2n(q+1)xp} \qquad [3.2.22]$$

Each column of $\mathbf{G}_{zi}$ is given by

$$\underline{\mathbf{G}}_{zi_l} \equiv \begin{bmatrix} \frac{\partial}{\partial\theta_l}[\mathbf{C}\,\underline{v}_1\,\underline{w}_1'\,x_o] \\ \vdots \\ \frac{\partial}{\partial\theta_l}[\mathbf{C}\,\underline{v}_n\,\underline{w}_n'\,x_o] \\ \frac{\partial\lambda_1}{\partial\theta_l}[\mathbf{C}\,\underline{v}_1\,\underline{w}_1'\,x_o] \\ \vdots \\ \frac{\partial\lambda_n}{\partial\theta_l}[\mathbf{C}\,\underline{v}_n\,\underline{w}_n'\,x_o] \end{bmatrix}_{2nx1} \qquad [3.2.23]$$

These vectors are dependent on both the initial conditions of the system and the structural characteristics of the state-space matrices.

Each column of $\mathbf{G}_{zs}$ is given by

$$\underline{G}_{zs_l} \equiv \begin{bmatrix} \frac{\partial}{\partial \theta_l}[\mathbf{C}\,\underline{v}_1\,\underline{w}_1'\,\mathbf{B}] \\ \bullet \\ \bullet \\ \bullet \\ \frac{\partial}{\partial \theta_l}[\mathbf{C}\,\underline{v}_n\,\underline{w}_n'\,\mathbf{B}] \\ \frac{\partial \lambda_1}{\partial \theta_l}[\mathbf{C}\,\underline{v}_1\,\underline{w}_1'\,\mathbf{B}] \\ \bullet \\ \bullet \\ \bullet \\ \frac{\partial \lambda_n}{\partial \theta_l}[\mathbf{C}\,\underline{v}_n\,\underline{w}_n'\,\mathbf{B}] \end{bmatrix}_{2nx1} \qquad [3.2.24]$$

This matrix is dependent on ***only*** the structural characteristics of the state-space matrices.

The matrices **E** and **F** have row vectors

$$\underline{e}_k \equiv \begin{bmatrix} e^{\lambda_1 t_k} \\ \bullet \\ \bullet \\ \bullet \\ e^{\lambda_n t_k} \\ t_k e^{\lambda_1 t_k} \\ \bullet \\ \bullet \\ \bullet \\ t_k e^{\lambda_n t_k} \end{bmatrix}_{2nx1} \qquad [3.2.25]$$

$$\underline{f}_k \equiv \begin{bmatrix} e^{\lambda_1 t_k} \int_0^{t_k} e^{-\lambda_1 \tau} \underline{u}(\tau) d\tau \\ \bullet \\ \bullet \\ \bullet \\ e^{\lambda_n t_k} \int_0^{t_k} e^{-\lambda_n \tau} \underline{u}(\tau) d\tau \\ t_k e^{\lambda_1 t_k} \int_0^{t_k} (t_k - \tau) e^{-\lambda_1 \tau} \underline{u}(\tau) d\tau \\ \bullet \\ \bullet \\ \bullet \\ t_k e^{\lambda_n t_k} \int_0^{t_k} (t_k - \tau) e^{-\lambda_n \tau} \underline{u}(\tau) d\tau \end{bmatrix}_{2nx1} \qquad [3.2.26]$$

Both **E** and **F** are time dependent matrices.

3.2.2.2 *Eigenvalue and Eigenvector Sensitivity*

The computations in forming [3.2.23 – 26] are straightforward except for the partial derivatives of the eigenvectors and eigenvalues with respect to each parameter. Closed form expressions for these partials were developed by Crossley and Porter (1969). These partials are dependent on only the system eigenvalues and eigenvectors. Unfortunately, Crossley and Porter's work only easily applies to cases when all the elements of **A** are linear in θ. Therefore, the eigenvalue and eigenvector sensitivities were determined by numerically perturbing each parameter in **A** and using the definition of a derivative to approximate the true sensitivity.

3.2.2.3 *Identifiability Condition*

To converge to the optimal parameter estimates for the linear state space model in [3.2.15 – 17] using the standard linearization approach requires that the least-squares problem

$$\underline{Y} = \mathbf{S}\Delta\underline{\theta} \qquad [3.2.27]$$

be solved at each step of the iteration. A necessary and sufficient condition for the solution is that **S** have rank p. More insight can be gained by examining the decomposition of **S** into its structural and time dependent parts

$$\mathbf{S} = [\mathbf{E}, \mathbf{F}]\mathbf{G} \qquad [3.2.6]$$

A necessary condition for **S** to have rank p is that **G** have rank p. Since **G** is dependent only on the system matrices **A** , **B** , **C** , and $\underline{x}_o$ at $\hat{\underline{\theta}}$, the requirement on the

rank of **G** becomes a structural condition on identifiability. Assuming all of the initial conditions are zero,

$$\mathbf{G}_{zI} = [\mathbf{0}], \qquad [3.2.28]$$

and therefore the requirement on the rank of $\mathbf{G}_{z_0}$ becomes the structural condition on identifiability.

Therefore, to implement Reid's method to determine a system's identifiability, the following algorithm should be used:

1. Formulate $\mathbf{G}_{z_0}$ (equation 3.2.24).
2. Determine the eigenvalue and eigenvector sensitivities from the **A** matrix.
3. Compute the rank of $\mathbf{G}_{z_0}$.

If the rank of $\mathbf{G}_{z_0}$ is greater than or equal to the number of estimated parameters, then the model is identifiable.

3.3 Conclusions

Grewal and Glover (1976) and Reid (1979) developed two distinct methods for determining identifiability of dynamic systems. Each method has several distinct advantages and disadvantages.

The state-space formulation of dynamic systems was used in the development of both methods. It offers the advantage of retaining physical significance to the parameters. Both identifiability methods are applicable to rail vehicle parameter estimation because they use the state space-method. Much rail vehicle parameter estimation has been performed using state-space dynamic modeling. Also, the

state-space method for rail vehicle modeling allows the determination of critical speeds and other stability issues which are essential in the evaluation of the safety aspects of lightweight rail vehicles.

One advantage of Grewal and Glover's technique is that the method can be extended to global identifiability as shown by Walter et al. (1979). Another advantage is the method permits an algebraic formulation of the identifiability condition which gives insight into the physical structure of the sensitivity matrix. This advantage, however, can also be the downfall of this method, especially for higher order models. Without the use of sophisticated algebraic manipulators for models even of low order, the algebra can become prohibitively complex. Norton (1980) also recognized this as a limitation of Grewal and Glover's method.

Another disadvantage of Grewal and Glover's technique is the method does not provide much quantitative information on the identifiability. With the algebraic formulation of even low order models, high order polynomials are formed and are sensitive to small changes in the parameter values. Therefore, in checking for the local identifiability of the model, small changes in the parameters produce large changes in the Jacobian of the Markov parameter matrix.

Reid's method uses a different approach from Grewal and Glover's method. Reid used a sensitivity matrix computed from reciprocal eigenvalues and eigenvectors to create the identifiability condition. Reid's method also allows an optimization of estimation accuracy through selection of experimental design variables such as input signal, sensor location, and sample interval. The estimation accuracy can be optimized by minimizing the condition number of the sensitivity matrix, **S**. With this information, an engineer can more easily plan identification experiments.

Reid's method does have several disadvantages, though. First, the method cannot be implemented algebraically. Also, the method is conceptually more complex and more difficult to implement than Grewal and Glover's method.

The advantages and disadvantages of the identifiability techniques are summarized in Table 2. Both methods are useful for determining the identifiability of rail vehicle models. Each method is developed differently and yields distinct insight into identifiability problems. Therefore, at least for a simple test case, both methods will be used.

In the next chapter, an example of the identifiability problem is given. Then, the two methods discussed in this chapter will be applied to the example problem and conclusions on the usefulness of the methods will be more adequately determined.

Table 2. Identifiability Methods

Method	*Advantages*	*Disadvantages*
Grewal and Glover	uses state-space formulation	difficult algebra for higher order models
	algebraic formulation of identifiability condition	high order polynomials affect sensitivity
	can be extended for global identifiability	no information on experimental design quality
	conceptually simple	
Reid	uses state-space formulation	algebraic formulation not possible
	can be used to optimize estimation accuracy	cannot be extended for global identifiability
	insight into experiment planning	conceptually complex
		difficult to implement

Chapter 4

Identifiability and Estimation Test Problem

4.1 Purpose

This chapter contains the development of a test problem for the parameter estimation of a single degree of freedom spring-mass-damper model. An output error method was used with frequency response functions as the data. This test problem was formulated to provide the following:

1. familiarization with the estimation technique
2. determination of parameter estimation results from nonidentifiable systems
3. an evaluation of the identifiability techniques
4. an evaluation of the data processing techniques

Also, all of the techniques were chosen to be extendable for the actual rail vehicle estimation.

4.2 Problem Formulation

Two different model structures were used for the test problem. Both of the model structures are one degree-of-freedom linear spring-mass-damper models. These represent subsets of standard rail vehicle models. The models differ in the location and type of input used to excite the system. Figure 2 shows these two models. Table 3 contains the parameter values used in both cases.

4.2.1 Displacement-Input Displacement-Output

The displacement-input displacement-output model in Figure 2 is represented by the differential equation

$$m\ddot{y} + c\dot{y} + ky = c\dot{z} + kz \qquad [4.2.1]$$

or

$$\ddot{y} + \left(\frac{c}{m}\right)\dot{y} + \left(\frac{k}{m}\right)y = \left(\frac{c}{m}\right)\dot{z} + \left(\frac{k}{m}\right)z \qquad [4.2.2]$$

where

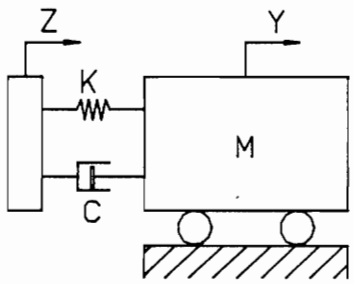
m = mass

c = linear damping constant

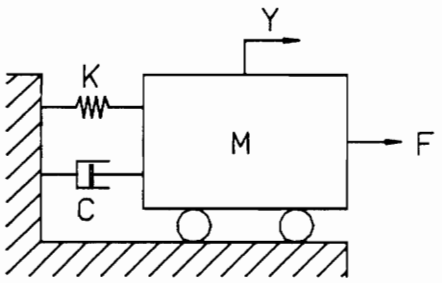
k = linear spring stiffness

and

z = input displacement



(A) DISPLACEMENT-INPUT DISPLACEMENT-OUTPUT



(B) FORCE-INPUT DISPLACEMENT-OUTPUT

Figure 2. Test Models: (a) displacement-in displacement-out (b) force-in displacement-out

Table 3. Parameter values used in model

Parameter	*Units*	*Value*
mass	kg	1.00
stiffness	$\frac{N}{m}$	25.0
damping	$\frac{N}{\left(\frac{m}{s}\right)}$	1.00

The state-space implementation of this equation is

$$\dot{\underline{x}} = \mathbf{A}\,\underline{x} + \mathbf{B}\,\underline{u} \qquad [4.2.3]$$

$$\underline{y} = \mathbf{C}\,\underline{x} + \mathbf{D}\,\underline{u} \qquad [4.2.4]$$

or, placing the values in the matrices,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\left(\frac{k}{m}\right) & -\left(\frac{c}{m}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} z \qquad [4.2.5]$$

$$y = \begin{bmatrix} \left(\frac{k}{m}\right) & \left(\frac{c}{m}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} z \qquad [4.2.6]$$

This model represents a subset of the rail vehicle tested. In the test condition, the rail vehicle sits on four actuators whose displacements excite the wheels of the vehicle. The primary signals measured on the rail vehicle are displacements. Therefore, the vehicle is an extension of the displacement input model. Figure 2 and Figure 3 show the analogy between the displacement input model and the actual rail vehicle.

The engineers at the Transportation Test Center formulated FRFs using input and output displacement measurements. Thus, an analysis of this model provides considerable insight into the rail vehicle estimation problem.

Examinination of equations [4.2.5] and [4.2.6] shows that an identifiability problem exists. In these equations the three parameters m , c, and k appear only in the combinations $\left(\frac{k}{m}\right)$ and $\left(\frac{c}{m}\right)$. Therefore, only the ratio of these parameters can be identified, not each individual parameter. A different approach formulates the equations in terms of natural frequency and damping ratio, instead of mass, damping, and stiffness. In this formulation only the parameters of natural frequency and damping ratio exist, so two parameters characterize the system. Clearly, three parameters cannot be identified.

4.2.2 Force-Input Displacement-Output

The force-input displacement-output model in Figure 2 is described by the differential equation

$$m\ddot{y} + c\dot{y} + ky = F \qquad [4.2.7]$$

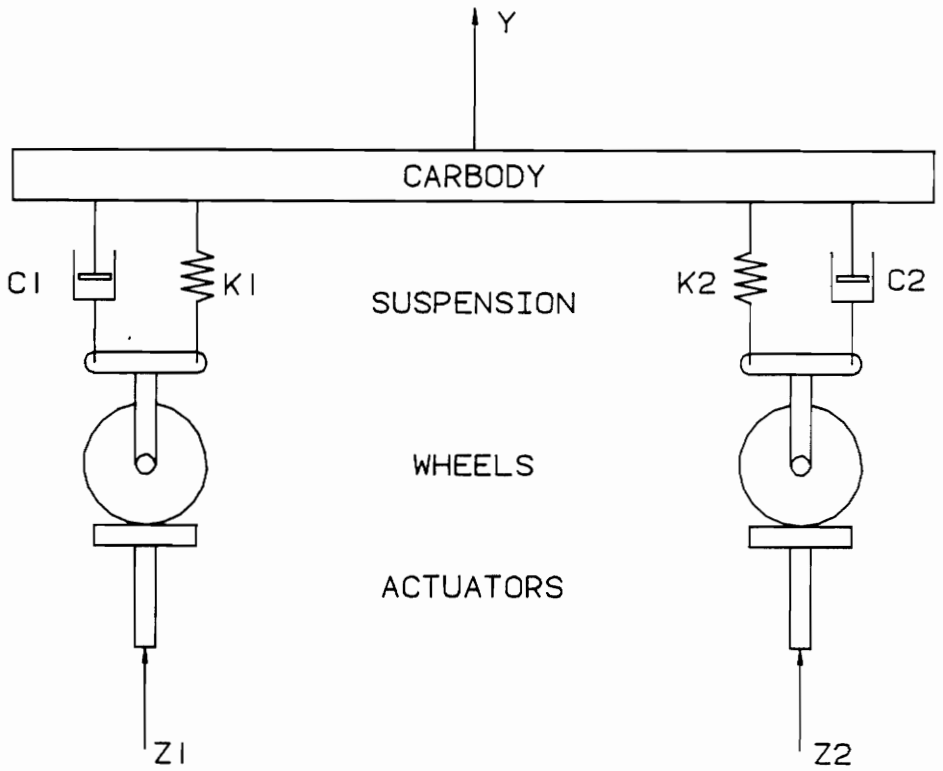


Figure 3. Basic rail vehicle similarities to displacement-input model

or

$$\ddot{y} + \left(\frac{c}{m}\right)\dot{y} + \left(\frac{k}{m}\right)y = \left(\frac{F}{m}\right) \qquad [4.2.8]$$

where the parameters are the same as in the previous case and F is the magnitude of the force applied to the system. The state space implementation of this equation is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\left(\frac{k}{m}\right) & -\left(\frac{c}{m}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \left(\frac{1}{m}\right) \end{bmatrix} F \qquad [4.2.9]$$

$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} F \qquad [4.2.10]$$

Equations [4.2.9] and [4.2.10] indicate that there are three parameter combinations that can be identified: $\frac{k}{m}$, $\frac{c}{m}$, and $\frac{1}{m}$. Therefore, the earlier problem with identifiability of the displacement-input model does not arise.

4.2.3 Forcing Function

The forcing function used for the test case described in this chapter was a linear sine sweep through the system's damped natural frequency. The forcing function used in the vehicle tests was a ramp-dwell sine sweep. It is similar to the function used in the test case. The forcing function in the test case is represented by the equation

$$F = F_o \sin[\{\omega_o + (\omega_1 t)\} t]. \qquad [4.2.11]$$

where F and F_o are either displacement or force values depending on the model used. The natural frequency of the system was 0.80 Hz and the damping ratio was 0.1. A frequency range of 0.1 to 2 Hz gave adequate system response for parameter estimation purposes.

4.2.4 Solution Method

The next step in the test case was simulating the model response to the forcing function. Two data records were required to perform the estimation, the input (displacement or force) and the output (displacement). These two records provided all of the data needed for the estimation procedure. The input and output records for a representative displacement-input model are shown in Figure 4.

The data was simulated by numerically integrating the model response using the Runge-Kutta-Fehlberg method developed by Forsythe et al. (1977). This method is frequently used for solving rail vehicle dynamics models (Fries and Coffey, 1987). The initial conditions of the model were set to zero. The time step used in the integration was 0.01 second and about 41,000 data pairs were calculated.

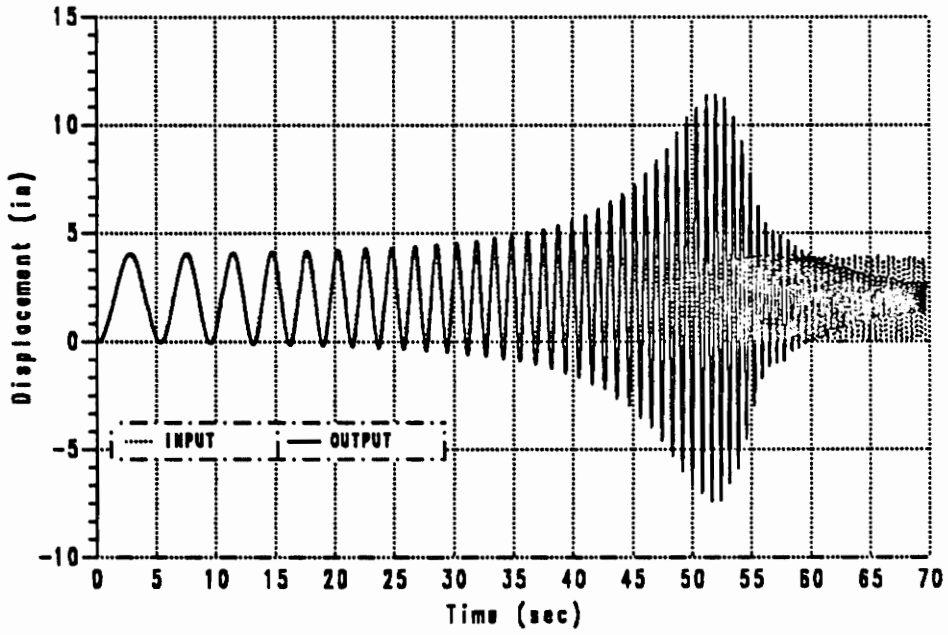


Figure 4. Comparison of input and output records

4.3 Data Processing

4.3.1 Transformation to Frequency Response Functions

Basic data processing techniques were used to transform the data from discrete time domain data to frequency response functions. The conversion from time domain to FRFs is summarized in six steps:

1. Transform functions to mean of zero by removing offset and trend.
2. Discard data further than 4 standard deviations from the mean.
3. Window the data to enforce periodicity.
4. Fast Fourier Transform (FFT) the data in groups of 2^n.
5. Compute the Power Spectral Densities (PSDs), Cross Spectral Densities (CSDs), and ensemble average the transformed data.
6. Compute the frequency response function.

All of these techniques are well documented and widely used in engineering work. Each of the steps is detailed below.

4.3.1.1 Transform Functions to Mean of Zero

The time response data was first processed by breaking it into sections of length 2^{12} (4096) points. Breaking the data into sections allowed relatively small amounts of data to be processed at a time and gave the frequency response functions good frequency resolution.

The next step in the data processing removed any drift in the instrumentation by performing a linear regression of the data. The data was fit to a straight line with a calculated mean and trend that were then subtracted from each data point. This procedure only minimally affected the simulated data but is necessary for most experimental data.

4.3.1.2 Discard Bad Data

After removing the offset and drift of the data records, the mean and variance of the residual data were computed. Then, the residual data was scanned for points exceeding four standard deviations from the mean. These points were replaced with a number equal to the mean value of the processed sequence. Again, this processing was of minimal use for the simulated data because the data was already well conditioned. However, it was valuable for the experimental data especially to replace data affected by some random disturbance during the tests.

4.3.1.3 Window Data

To transform the data into the frequency domain, standard Fast Fourier Transform (FFT) techniques were used. These techniques require a periodic data set. If the data set is not periodic then spectral leakage distorts the PSDs, ultimately giving inaccurate FRFs. To reduce the distortion, a windowing algorithm is needed. References and algorithms for data windowing are found in Bendat and Piersol (1986), Press et al. (1986), and Harris (1978).

Many windowing algorithms exist. Four of the most popular windowing techniques, boxcar window (no window), Parzen window, Welch window, and Hanning window are shown in Figure 5. The differences among the Parzen, Welch, and Hanning windows are subtle, and they affect the narrowness of the spectral leakage functions. Press et al. (1986) recommended using either the Parzen window or the Welch window. Thus, the simple Parzen window was used in this estimation work. Figure 6 shows the effect of the Parzen window on typical data.

4.3.1.4 Fast Fourier Transform Data

After the preliminary processing on the time data, the data sets were transformed into the frequency domain. A standard Fast Fourier Transform (FFT) algorithm was used to process the data. Two references for standard FFT methods are Bendat and Piersol (1986) and Press et al. (1986).

The transformation into the frequency domain is represented here by changing from lower case to upper case characters. Therefore, the transformations computed for the input and output sequences are:

$x(t) \rightarrow X(j\omega)$

$y(t) \rightarrow Y(j\omega)$

The frequency resolution of the FFT depends on both the sample rate of the data and the record length (number of samples per record). The relationship between these variables is

$$\Delta f = \frac{f_s}{N} \qquad [4.3.1]$$

where

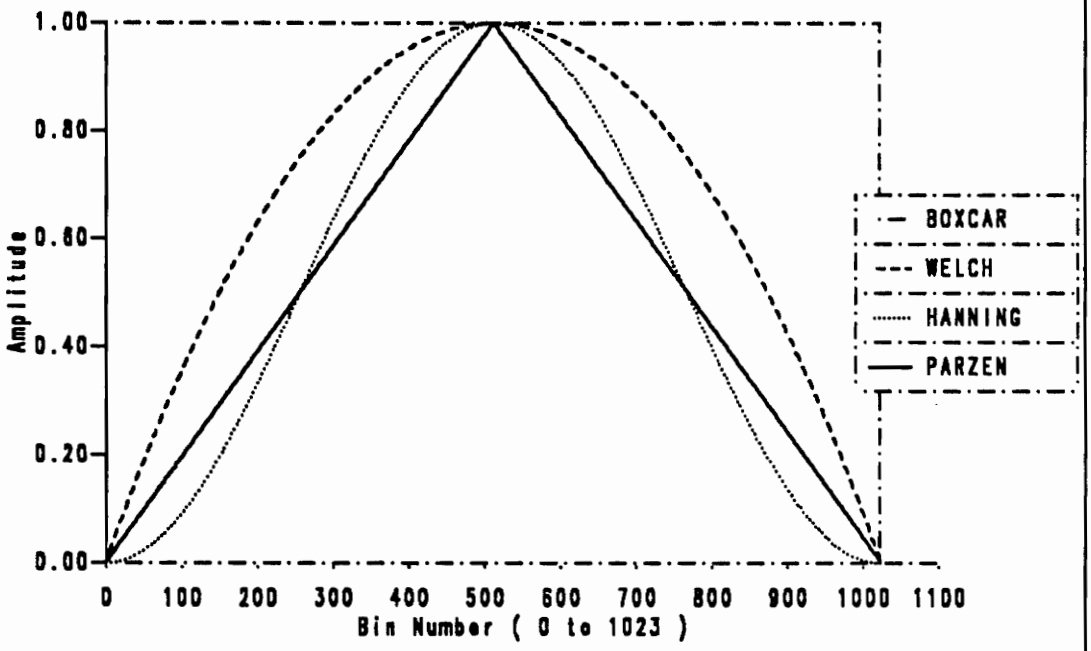


Figure 5. Common data windows (Press et al., 1986)

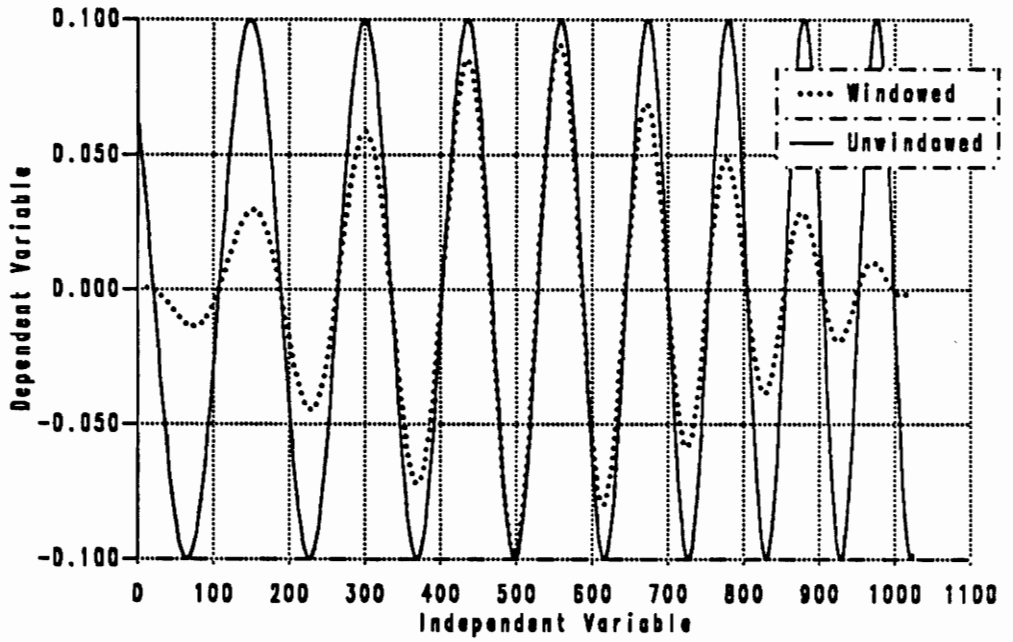


Figure 6. Effect of data windowing

Δf = frequency resolution

f_s = sample frequency

and

N = record length.

Frequency resolution is an important factor in analyzing dynamic systems, especially when the system is very lightly damped. When the system is lightly damped, the width of the FRF peak decreases. Therefore, if the frequency resolution of the transform is too large, the peak amplitude might be computed incorrectly.

4.3.1.5 Compute the PSDs and Ensemble Average the Data

The next step in the processing is to compute the power spectral densities of each transformed data record and ensemble average the records. The ensemble averaging incorporates the information from each section of data, here about 10, into one record. The ensembled power spectral densities (PSDs) and cross spectral densities (CSDs) are computed by the following equations:

$$G_{xx}(j\omega) = \frac{1}{n_{ds}} \sum_{i=1}^{n_{ds}} X_i^*(j\omega) X_i(j\omega) \qquad [4.3.2]$$

$$G_{xy}(j\omega) = \frac{1}{n_{ds}} \sum_{i=1}^{n_{ds}} X_i^*(j\omega) Y_i(j\omega) \qquad [4.3.3]$$

$$G_{yy}(j\omega) = \frac{1}{n_{ds}} \sum_{i=1}^{n_{ds}} X_i^*(j\omega) Y_i(j\omega) \qquad [4.3.4]$$

where

G_{xx} = power spectral density of input

G_{xy} = cross spectral density of input and output

G_{yy} = power spectral density of output

and

n_{ds} = number of data sections.

4.3.1.6 Compute the Transfer Function

The final step of data processing is to compute the frequency response function. Bendat and Piersol (1986), Rost and Leuridian (1985), and Mitchell (1982) all compared several different methods to compute frequency response functions. Mitchell (1982) proposed three distinct methods: $\tilde{H}_1$, $\tilde{H}_2$, and $\tilde{H}_3$. The methods are defined as follows:

$$\tilde{H}_1(f) = \frac{G_{xy}}{G_{xx}} \qquad [4.3.5]$$

$$\tilde{H}_2(f) = \frac{G_{yy}}{G_{xy}^*} \qquad [4.3.6]$$

and

$$\tilde{H}_3(f) = \frac{\tilde{H}_1(f) + \tilde{H}_2(f)}{2} \qquad [4.3.7]$$

These FRFs contain slightly different noise immunity characteristics. To examine these characteristics, the following definitions are needed:

H_o = true frequency response function

G_{nn} = power spectral density of the input random measurement noise

G_{mm} = power spectral density of the output random measurement noise

G_{uu} = true power spectral density of the input

G_{vv} = true power spectral density of the output

and

G_{uv} = cross-spectral density between the true input and true output

Figure 7 shows the block diagram of a single-input single-output system with noise.

The relationship of the frequency response functions $\tilde{H}_1$, $\tilde{H}_2$, and $\tilde{H}_3$ to random errors and the true FRFs are given by Mitchell as:

$$\tilde{H}_1(f) = \frac{H_o(f)}{1 + \frac{G_{nn}}{G_{uu}}} \qquad [4.3.8]$$

$$\tilde{H}_2(f) = H_o(f)\left[1 + \frac{G_{mm}}{G_{vv}}\right] \qquad [4.3.9]$$

and

$$\tilde{H}_3(f) = \frac{H_o(f)}{2}\left[\frac{1}{\left(1+\frac{G_{nn}}{G_{uu}}\right)} + \left(1+\frac{G_{mm}}{G_{vv}}\right)\right]. \qquad [4.3.10]$$

From these estimators, several conclusions can be made:

1. $\tilde{H}_1$ is contaminated by input measurement noise.
2. $\tilde{H}_2$ yields an estimate greater than H_o and is sensitive to output measurement noise. Because there is no input measurement error, this estimator provides good estimates near the natural frequencies

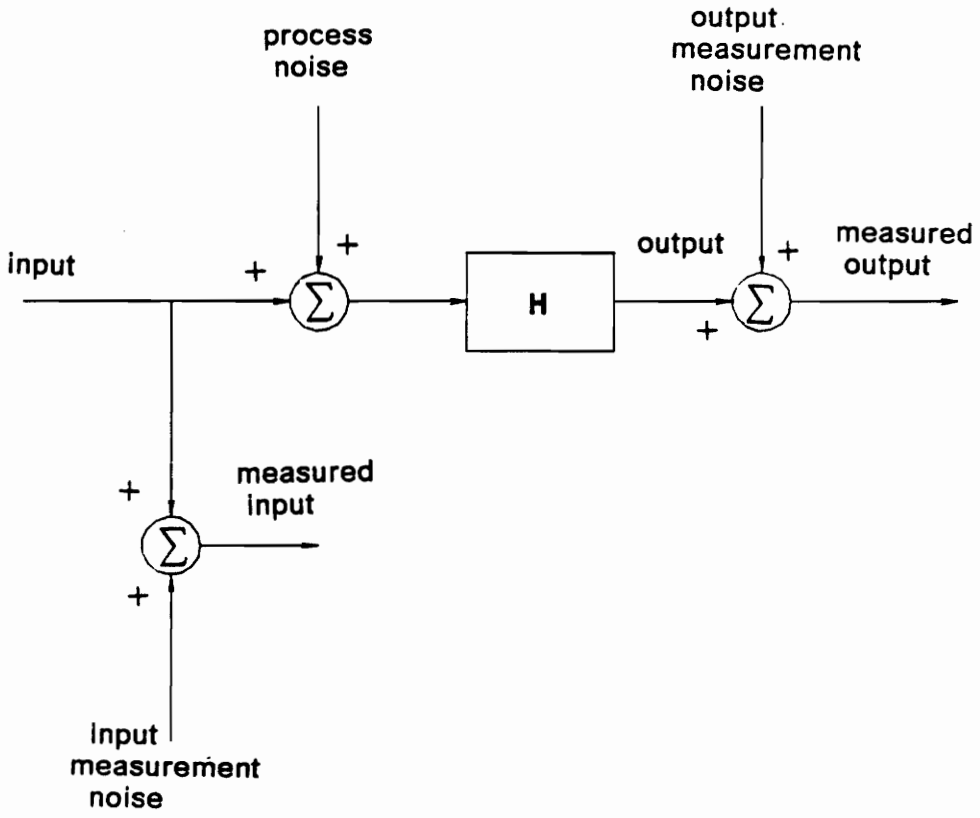


Figure 7. Single-input single-output system with noise

of the system.

3. $\tilde{H}_3$ is the average of $\tilde{H}_1$ and $\tilde{H}_2$. Therefore, $\tilde{H}_3$ provides a good general estimate for most experiments unless the FRF around resonance is very important.

For this work, $\tilde{H}_1$ was chosen for the test cases because there was no measurement error. For the vehicle estimation case, however, $\tilde{H}_2$ was used because the primary information for the estimation was near the natural frequency. The computed frequency response functions for the two test cases are shown in Figure 8 and Figure 9.

4.3.2 Parameter Estimation of Simulated Data

Once the frequency response functions of the data were computed, the parameter estimation was performed. The output-error estimation method developed in Chapter 2 was used with a few alterations. Because the data was simulated, no measurement error information was available. Therefore, the observation error covariance matrix, $\mathbf{S}_{zz}$, was simply the identity matrix. Also, since all of the parameters could easily be identified with one observation set, the parameter error covariance matrix, $\mathbf{S}_{rr}$, was not needed. The parameter error covariance matrix was also not originally used to influence the relative changes in the parameters discussed in section [2.3.2.2] . Later, $\mathbf{S}_{rr}$ was used to determine its effect on the estimation of a nonidentifiable system. These changes originally transformed the parameter estimation to the simple unweighted least-squares class. Thus, the function minimized was

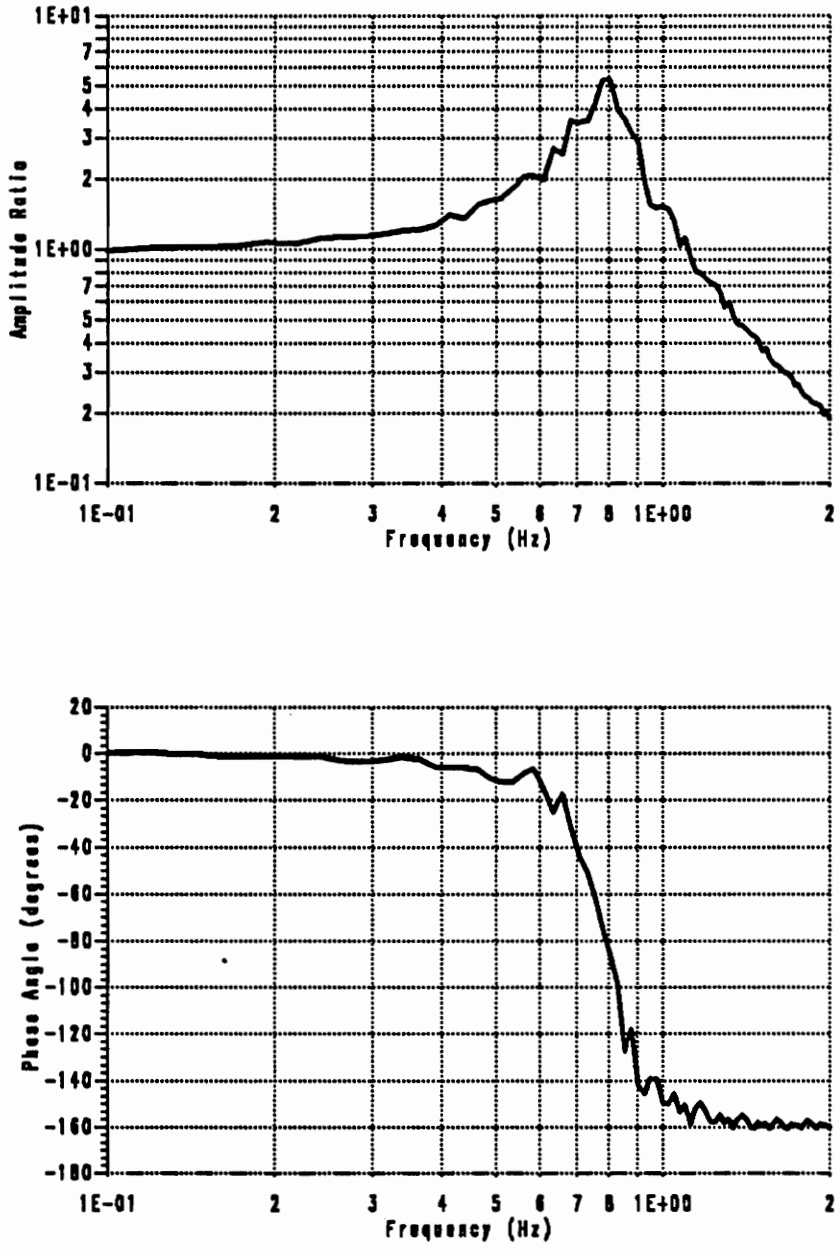


Figure 8. Computed FRF of displacement-input displacement-output model

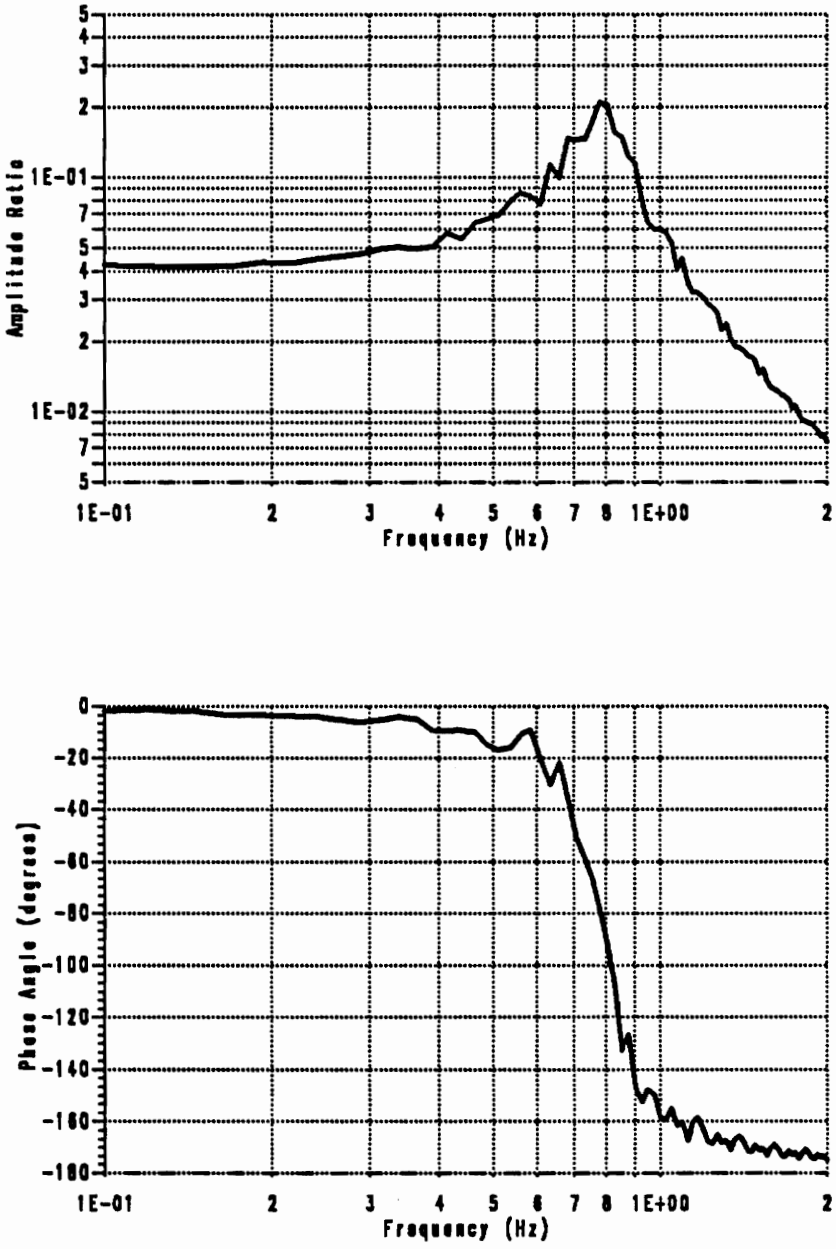


Figure 9. Computed FRF of force-input displacement-output model

$$J = (\hat{\underline{H}} - \underline{H})' (\hat{\underline{H}} - \underline{H}) \qquad [4.3.11]$$

where

$\hat{\underline{H}}$ = frequency response function computed from measured outputs

$\underline{H}$ = model frequency response function

= $\mathbf{C}(j\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$

The FRFs covered the frequency range from 0.1 Hz to 2 Hz. The frequency resolution of the data was 0.0244 Hz.

4.3.3 Results

The parameter estimation program estimated the parameters m , c, and k from the frequency response function magnitudes only. Figure 10 and Figure 11 show the convergence of the FRF magnitudes to the experimental data. Figure 12 and Figure 13 show the performance index convergence for each model case. The performance indices are not the same magnitude because of differences in the two model structures. The FRF magnitude at low frequencies for the force-input model approaches $\frac{1}{k}$, whereas the magnitude for the displacement-input model approaches one. Therefore the performance indices can be more easily compared with a scale factor of k^2 applied to the force-input index.

Figures 10-13 indicate good convergence and a successful estimation of the model parameters. However, when the actual values of the estimated parameters are examined, an unexpected result appears: the estimated values of the parameters are widely different. The initial values of the parameters and the estimates of the parameters after the performance index converged (5 iterations) are shown in

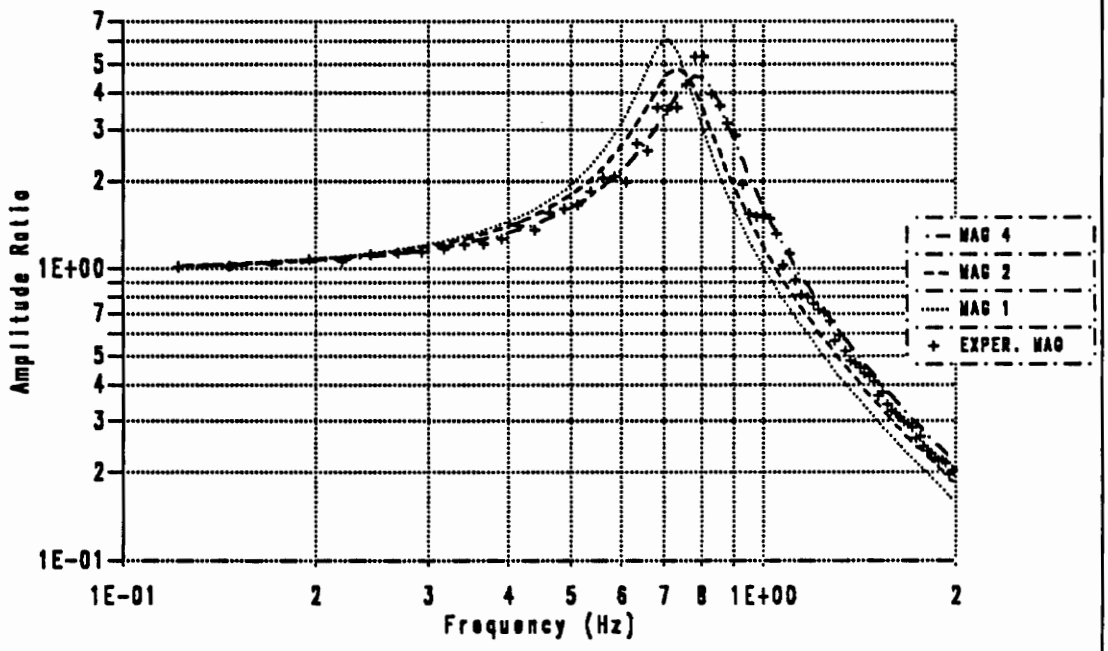


Figure 10. Magnitude convergence for the displacement-input model

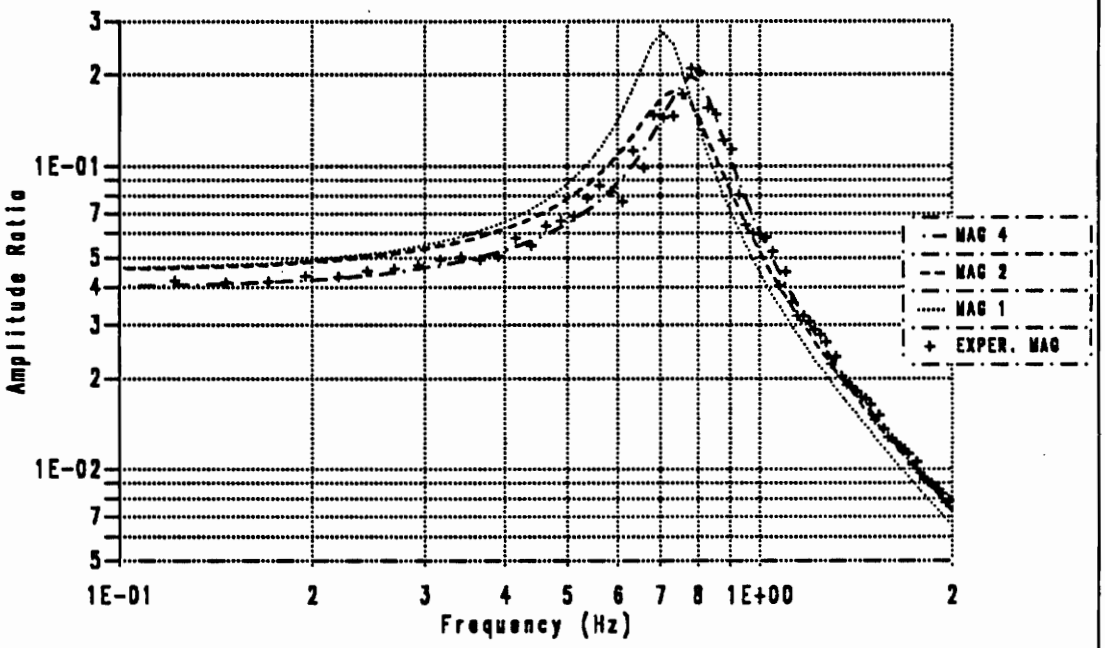


Figure 11. Magnitude convergence for the force-input model

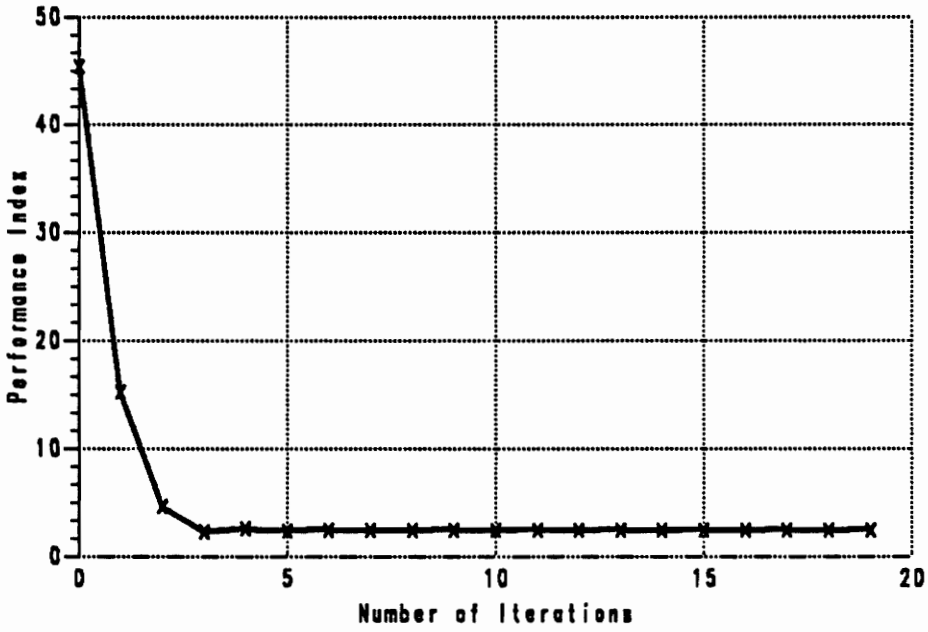


Figure 12. Convergence of performance index for the displacement-input model

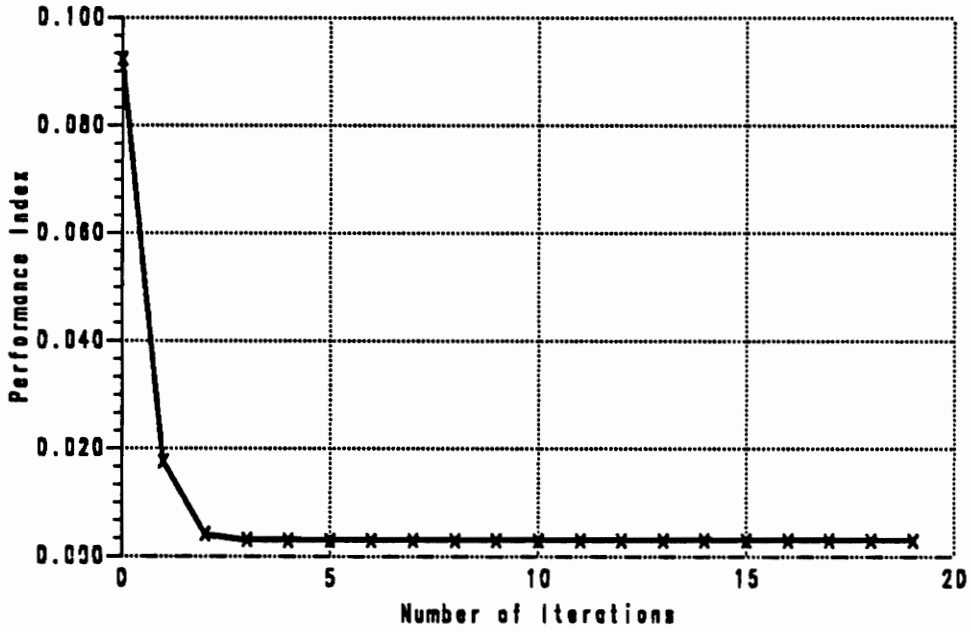


Figure 13. Convergence of performance index for the force-input model

Table 4. The percentage errors of each identified parameter to the true value at this point are shown in Table 5.

The convergence of each individual parameter value is seen in Figures 14-16. These figures show the convergence of the force-input model to stable, nonzero values within four iterations. For the displacement-input case, however, the parameters continue to change even though the performance index converges.

The displacement-input model parameters converge to zero because a least-squares estimator without any Bayesian term was used. The addition of the Bayesian term forces the estimation to converge to a nonzero value, thus making the identifiability problem even more difficult to detect. Figure 17 thru Figure 20 show the Bayesian term effects the convergence properties of the nonidentifiable system. These figures show how the Bayesian term effectively hides the identifiability problem within the parameter estimation.

The estimation of the displacement-input model also is suspect from the phase angle plots. Phase angle data is not included in the performance index of either model and therefore provides a check on the reliability of the model. Figure 21 and Figure 22 show the convergence of the phase angle data for the two test cases. For the force-input model, the phase angle is modeled more accurately at each iteration whereas for the displacement-input model, the phase angle data does not converge adequately at higher frequencies. This divergence indicates a poor estimation.

These results indicate the difficulty in identifying the parameters of certain system structures. For the displacement-input system, the identifiability problem is evident from the first formulation of the problem where only two combinations of the parameters, $\left(\frac{c}{m}\right)$ and $\left(\frac{k}{m}\right)$, appeared in the differential equation. However, when

Table 4. Comparison of estimated parameter values

Parameter	*True Value*	*Initial Guess*	*Estimated value for d-in d-out model*	*Estimated value for f-in d-out model*
mass	1.00	1.10	0.641	1.002
stiffness	25.0	22.0	16.1	24.9
damping	1.00	0.800	0.720	1.02

the model becomes more complex than this simple example, the identifiability issues are not easily evident.

4.3.4 Implications of the Identifiability Problem

These examples show how nonidentifiable systems affect parameter estimation results. Parameter estimation results of even low order models can easily mislead an investigator, especially if he does not carefully examine the model structure or scrutinize all of the estimation results. With models of large order,

Table 5. Comparison of estimated parameter errors

Parameter	*True Value*	*Percent error for d-in d-out model*	*Percent error for f-in d-out model*
mass	1.00	36	0.20
stiffness	25.0	35	.40
damping	1.00	28	2.0

identifiability problems can be very difficult to determine simply from looking at the model structure.

In this example, the identifiability problems are not evident until a close examination of the estimation results is performed. This hides the identifiability problem, and can result in other factors such as measurement noise or process noise being blamed for the error. Measurement and process noise certainly do affect the estimation results, but they change the results independently of the identifiability of the system.

Several common themes permeate the parameter estimation literature. They will be examined using the results from the identifiability test cases. The themes are:

1. Parameters obtained are sometimes not reasonable.

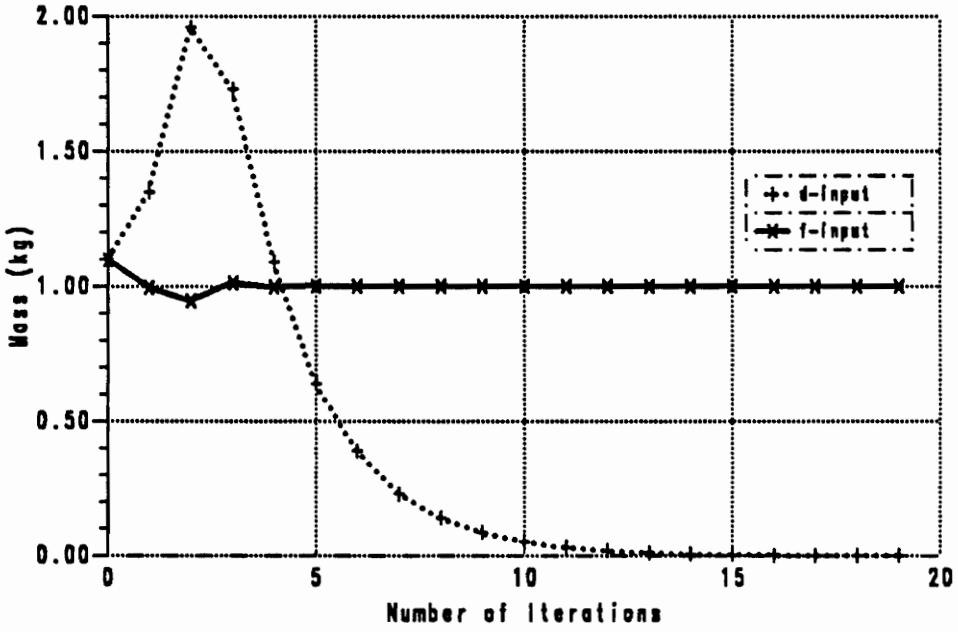


Figure 14. Convergence of the mass parameter for both models

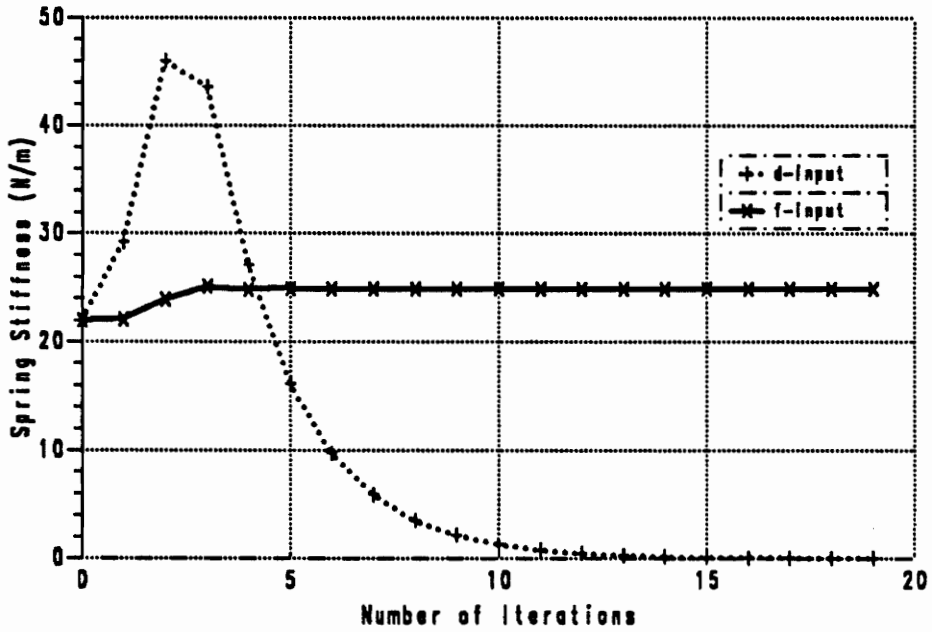


Figure 15. Convergence of the stiffness parameter for both models

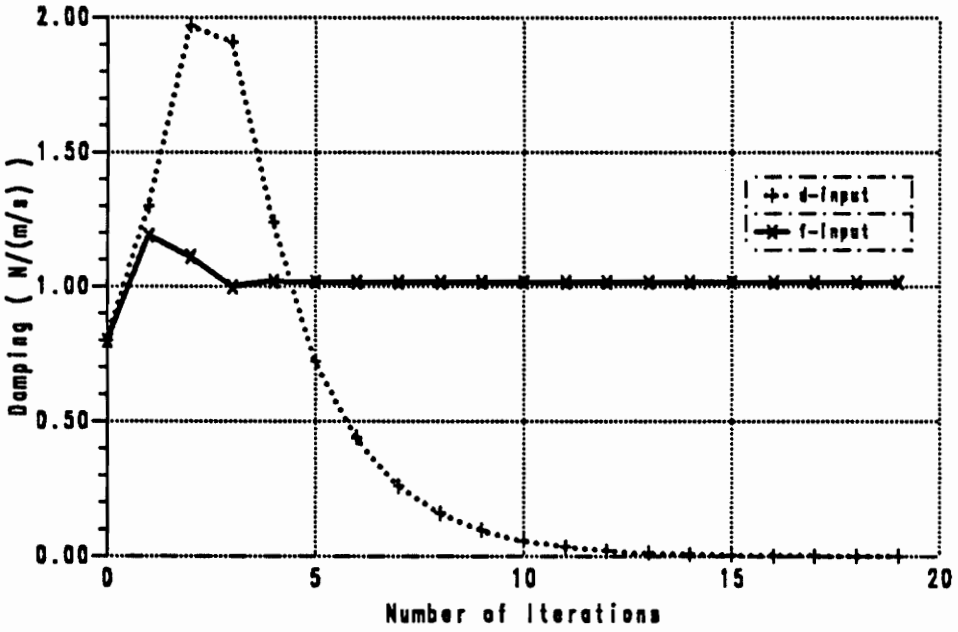


Figure 16. Convergence of the damping parameter for both models

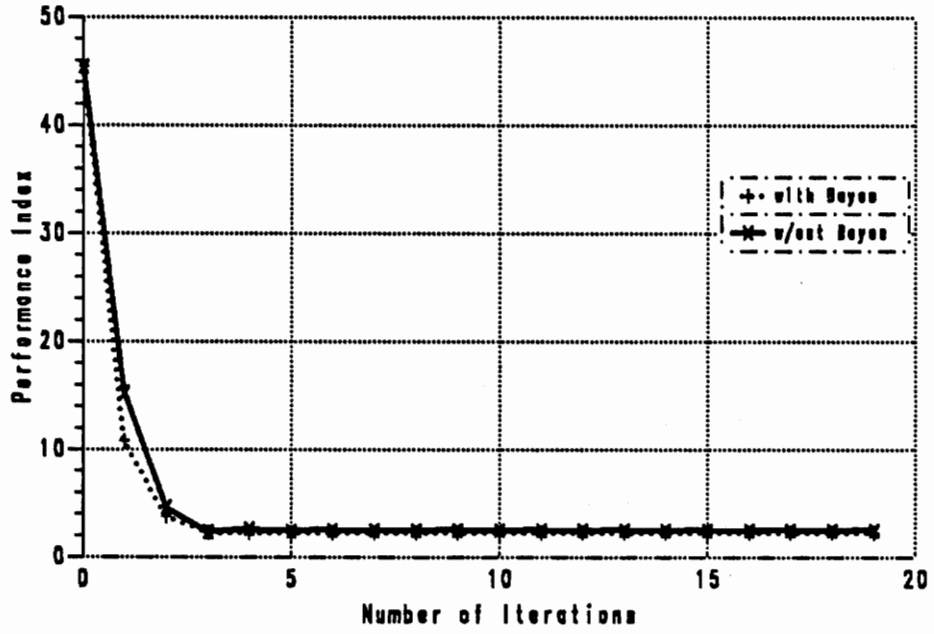


Figure 17. Convergence of the performance index with and without the Bayes term

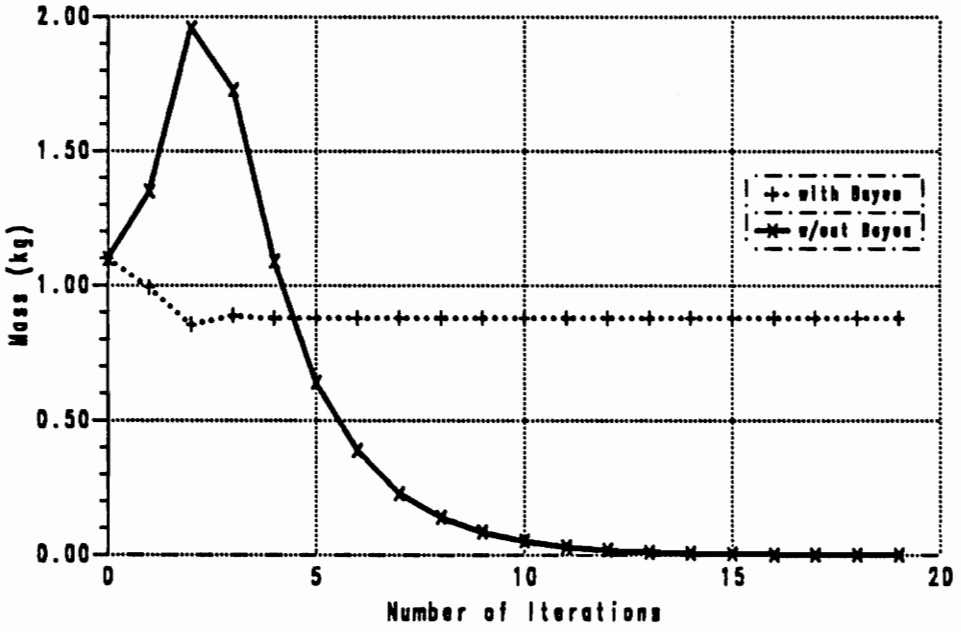


Figure 18. Convergence of the mass parameter with and without the Bayes term

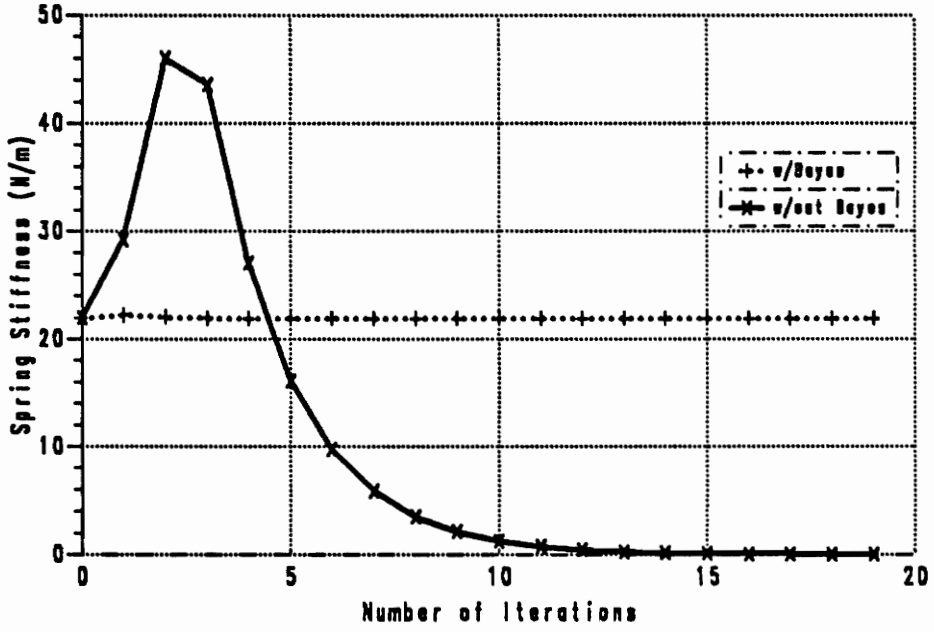


Figure 19. Convergence of the stiffness parameter with and without the Bayes term

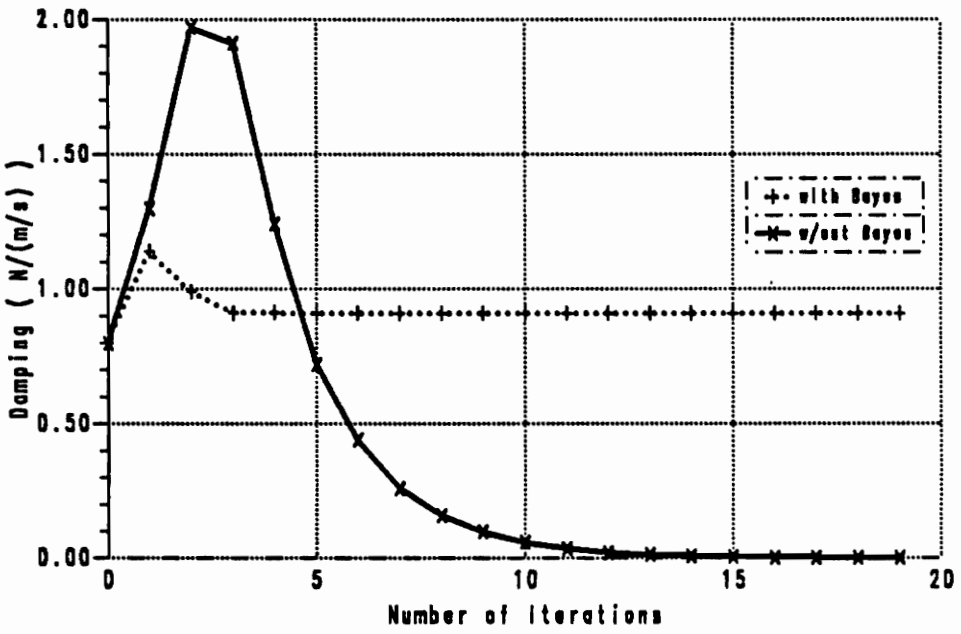


Figure 20. Convergence of the damping parameter with and without the Bayes term

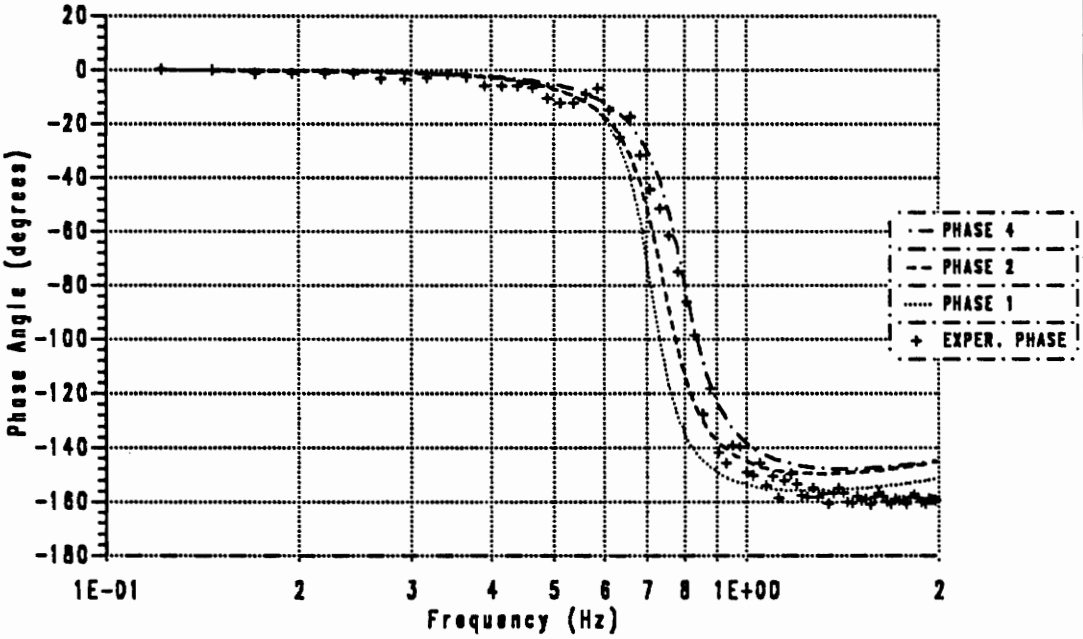


Figure 21. Phase angle convergence for the displacement-input model

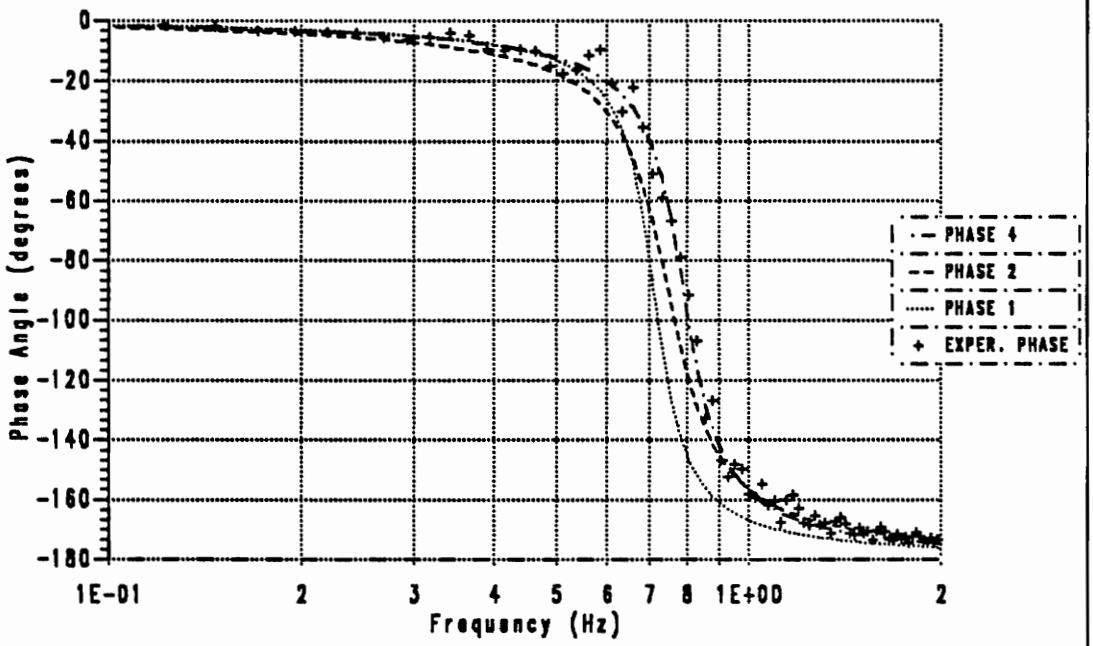


Figure 22. Phase angle convergence for the force-input model

2. Parameters may be changed without affecting the performance index
3. A model may be adjusted to match certain data without actually improving the model

Each of these statements is easily applicable to the estimation of the displacement-input case.

Parameters obtained are not reasonable: This result is obvious from the test model. In the test model, the true values were used to create the simulation. Table 4 shows the parameter values obtained for the displacement-input model after five iterations were incorrect by at least 28 percent.

Parameters may be changed without affecting the performance index: Figures 9 and 10 show the performance index converges within five iterations. However, Figures 11-13 show the individual mass, damping, and stiffness values continue to change even after the convergence of the performance index

A model may be adjusted to match certain data without actually improving the model: This result is actually a combination of the two previous results. Figure 10 shows the convergence of the magnitude estimation to the experimental data, even though the parameter values are still changing.

These three difficulties with the model structure are the result of the nonidentifiability of the parameters in the displacement-input model. The problem with identifying the parameters is seen quite easily in retrospect. However the difficulty was not obvious at the beginning of the modeling. These same difficulties are more frustrating for those working with larger models. The identifiability problems can also nullify the results of much experimentation and therefore increase testing costs. If the identifiability procedures developed by Grewal and Glover (1976) and Reid (1979) provide a priori information from the model structure, many

identifiability problems can be resolved before experimentation. Thus, in the next section, the two identifiability tests are applied to both the force-input and displacement-input models.

4.4 Identifiability Procedures Performed on Test Problem

Both Grewal and Glover's method and Reid's method are applied to the displacement-input and force-input models. The results from this development are used to form general conclusions about the utility of these techniques for more complex models and for the specific rail vehicle model identified in this thesis.

4.4.1 Force-Input Displacement-Output Model

4.4.1.1 Grewal and Glover's Method

The state space implementation of the force-input model is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\left(\frac{k}{m}\right) & -\left(\frac{c}{m}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \left(\frac{1}{m}\right) \end{bmatrix} F \qquad [4.4.1]$$

$$y = [1 \ 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \qquad [4.4.2]$$

where the parameters are defined in Section 2 of this chapter.

Grewal and Glover's method checks whether the rank of the Jacobian of the Markov parameter matrix is equal to the number of identified parameters in the system. The Markov parameter matrix from equation [3.2.13] is

$$\mathbf{G} \equiv \begin{bmatrix} \mathbf{CB} \\ \mathbf{CAB} \\ \mathbf{CA^2B} \\ \mathbf{CA^3B} \end{bmatrix} \qquad [4.4.3]$$

The individual terms of the matrix are

$$\mathbf{CB} = [1 \;\; 0]\begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} = [0] \qquad [4.4.4]$$

$$\mathbf{C[AB]} = [1 \;\; 0]\begin{bmatrix} \frac{1}{m} \\ -\frac{c}{m^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{m} \end{bmatrix} \qquad [4.4.5]$$

$$\mathbf{C[A^2B]} = [1 \;\; 0]\begin{bmatrix} -\frac{c}{m^2} \\ -\frac{k}{m^2} + \frac{c^2}{m^3} \end{bmatrix} = \begin{bmatrix} -\frac{c}{m^2} \end{bmatrix} \qquad [4.4.6]$$

$$\mathbf{C[A^3B]} = [1 \;\; 0]\begin{bmatrix} -\frac{k}{m^2} + \frac{c^2}{m^3} \\ \frac{kc}{m} + \frac{c^2 + kc + k^2}{m^3} - \frac{c^3 + kc^2}{m^4} \end{bmatrix} = \begin{bmatrix} -\frac{k}{m^2} + \frac{c^2}{m^3} \end{bmatrix} \qquad [4.4.7]$$

Therefore, the Markov parameter matrix is

$$\mathbf{G} = \begin{bmatrix} 0 \\ \frac{1}{m} \\ -\frac{c}{m^2} \\ -\frac{k}{m^2} + \frac{c^2}{m^3} \end{bmatrix}. \qquad [4.4.8]$$

The Jacobian of this matrix is

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{G}}{\partial m} & \frac{\partial \mathbf{G}}{\partial c} & \frac{\partial \mathbf{G}}{\partial k} \end{bmatrix} = \begin{bmatrix} -\frac{1}{m^2} & 0 & 0 \\ \frac{2c}{m^3} & -\frac{1}{m^2} & 0 \\ \frac{2k}{m^3} - \frac{3c^2}{m^4} & \frac{kc}{m^3} & -\frac{1}{m^2} \end{bmatrix} \qquad [4.4.9]$$

The first row of **J** was discarded because it contained only zeros and therefore did not add any information to the matrix. The determinant of the Jacobian matrix is equal to

$$\left[-\frac{1}{m^2} \right]^3 \qquad [4.4.10]$$

for any values of the parameters. Therefore, this matrix has rank equal to three. This result shows that the force-input case is identifiable for the three parameters m , c, and k.

4.4.1.2 *Reid's Method*

Reid's method of identifiability uses a decomposition of the solution to the state-space model. If each parameter in the state-space system is identifiable, then the rank of the matrix $\mathbf{G}_{zs}$ must be equal to the number of parameters. Each column of this matrix, derived in section [3.2.2.1], is

$$\underline{G}_{zs_i} \equiv \begin{bmatrix} \frac{\partial}{\partial \theta_i}[\mathbf{C}\,\underline{v}_1\,\underline{w}'_1\,\mathbf{B}] \\ \vdots \\ \frac{\partial}{\partial \theta_i}[\mathbf{C}\,\underline{v}_n\,\underline{w}'_n\,\mathbf{B}] \\ \left(\frac{\partial \lambda_1}{\partial \theta_i}\right)[\mathbf{C}\,\underline{v}_1\,\underline{w}'_1\,\mathbf{B}] \\ \vdots \\ \left(\frac{\partial \lambda_n}{\partial \theta_i}\right)[\mathbf{C}\,\underline{v}_n\,\underline{w}'_n\,\mathbf{B}] \end{bmatrix}_{2nx1} \qquad [4.4.11]$$

where

θ_i = i^{th} parameter ($\theta_1 = k$, $\theta_2 = c$, $\theta_3 = m$)

$\underline{v}_i$ = i^{th} right eigenvector of **A**

$\underline{w}_i$ = i^{th} left eigenvector of **A**

λ_i = i^{th} eigenvalue of **A**

The full expansion of the i^{th} column of this matrix, valid for both force-input and the displacement-input, is

$$\underline{G}_{zs_i} = \begin{bmatrix} \frac{\partial \mathbf{C}}{\partial \theta_i} \underline{v}_1 \underline{w}'_1 \mathbf{B} + \mathbf{C} \frac{\partial \underline{v}_1}{\partial \theta_i} \underline{w}'_1 \mathbf{B} + \mathbf{C} \underline{v}_1 \frac{\partial \underline{w}'_1}{\partial \theta_i} \mathbf{B} + \mathbf{C} \underline{v}_1 \underline{w}'_1 \frac{\partial \mathbf{B}}{\partial \theta_i} \\ \frac{\partial \mathbf{C}}{\partial \theta_i} \underline{v}_2 \underline{w}'_2 \mathbf{B} + \mathbf{C} \frac{\partial \underline{v}_2}{\partial \theta_i} \underline{w}'_2 \mathbf{B} + \mathbf{C} \underline{v}_2 \frac{\partial \underline{w}'_2}{\partial \theta_i} \mathbf{B} + \mathbf{C} \underline{v}_2 \underline{w}'_2 \frac{\partial \mathbf{B}}{\partial \theta_i} \\ \frac{\partial \lambda_1}{\partial \theta_i} \mathbf{C} \underline{v}_1 \underline{w}'_1 \mathbf{B} \\ \frac{\partial \lambda_2}{\partial \theta_i} \mathbf{C} \underline{v}_2 \underline{w}'_2 \mathbf{B} \end{bmatrix} \qquad [4.4.12]$$

For the force-input case, the partial derivatives of the matrices **B** and **C** are

$$\frac{\partial \mathbf{B}}{\partial \theta_1} = \frac{\partial}{\partial k} \begin{bmatrix} 0 \\ \left(\frac{1}{m}\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad [4.4.13]$$

$$\frac{\partial \mathbf{B}}{\partial \theta_2} = \frac{\partial}{\partial c} \begin{bmatrix} 0 \\ \left(\frac{1}{m}\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad [4.4.14]$$

$$\frac{\partial \mathbf{B}}{\partial \theta_3} = \frac{\partial}{\partial m} \begin{bmatrix} 0 \\ \left(\frac{1}{m}\right) \end{bmatrix} = \begin{bmatrix} 0 \\ -\left(\frac{1}{m^2}\right) \end{bmatrix} \qquad [4.4.15]$$

$$\frac{\partial \mathbf{C}}{\partial \theta_1} = \frac{\partial}{\partial k} [1 \;\; 0] = [0 \;\; 0] \qquad [4.4.16]$$

$$\frac{\partial \mathbf{C}}{\partial \theta_2} = \frac{\partial}{\partial c} [1 \;\; 0] = [0 \;\; 0] \qquad [4.4.17]$$

$$\frac{\partial \mathbf{C}}{\partial \theta_3} = \frac{\partial}{\partial \mathrm{m}}[1\ 0] = [0\ 0] \qquad [4.4.18]$$

The eigenvalue and eigenvector sensitivities were determined by perturbing each parameter in **A** and using the definition of a derivative to approximate the true sensitivity. For example, $\frac{\partial \lambda_1}{\partial \theta_1}$ was determined from

$$\frac{\partial \lambda_1}{\partial \theta_1} = \frac{\lambda_1^{\dagger} - \lambda_1}{\Delta \theta_1} \qquad [4.4.19]$$

where

λ_1 = eigenvalue of **A** matrix computed using θ_1

$\lambda_1^{\dagger}$ = eigenvalue of **A** matrix computed using $\theta_1 + \Delta\theta_1$

$\Delta\theta_1$ = small change in θ_1

Once the matrix was assembled, the rank of the matrix was determined using Matlab, a commercial software package. (Matlab, 1987) The technique used by Matlab is discussed in more detail in the next section.

For all combinations of values checked in the region near the initial values of the estimation, the rank of $\mathbf{G}_{zs}$ was three. Therefore, Reid's method indicates that the force-input model is identifiable for all three parameters m, c, and k.

4.4.2 Displacement-Input Displacement-Output Model

4.4.2.1 Grewal and Glover's Method

The state space implementation of the displacement-input model is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\left(\frac{k}{m}\right) & -\left(\frac{c}{m}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} z \qquad [4.4.20]$$

$$y = \begin{bmatrix} \left(\frac{k}{m}\right) & \left(\frac{c}{m}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad [4.4.21]$$

where the parameters are defined in Section 2 of this chapter.

For the displacement input case, the individual terms of the Markov parameter matrix are

$$\mathbf{CB} = \begin{bmatrix} \frac{k}{m} & \frac{c}{m} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{c}{m} \qquad [4.4.22]$$

$$\mathbf{C[AB]} = \begin{bmatrix} \frac{k}{m} & \frac{c}{m} \end{bmatrix} \begin{bmatrix} 1 \\ -\frac{c}{m} \end{bmatrix} = \frac{k}{m} - \frac{c^2}{m^2} \qquad [4.4.23]$$

$$\mathbf{C[A^2B]} = \begin{bmatrix} \frac{k}{m} & \frac{c}{m} \end{bmatrix} \begin{bmatrix} -\frac{c}{m} \\ -\frac{k}{m} + \frac{c^2}{m^2} \end{bmatrix} = -\frac{2kc}{m^2} + \frac{c^3}{m^3} \qquad [4.4.24]$$

$$\mathbf{C}[\mathbf{A}^3\mathbf{B}] = \begin{bmatrix} \frac{k}{m} & \frac{c}{m} \end{bmatrix} \begin{bmatrix} -\frac{k}{m} + \frac{c^2}{m^2} \\ \frac{2kc}{m^2} - \frac{c^3}{m^3} \end{bmatrix} = -\frac{k^2}{m^2} + \frac{3c^2k}{m^3} - \frac{c^4}{m^4} \qquad [4.4.25]$$

Therefore, the Markov parameter matrix is

$$\mathbf{G} = \begin{bmatrix} \frac{c}{m} \\ \frac{k}{m} - \frac{c^2}{m^2} \\ -\frac{2kc}{m^2} + \frac{c^3}{m^3} \\ -\frac{k^2}{m^2} + \frac{3kc^2}{m^3} - \frac{c^4}{m^4} \end{bmatrix}. \qquad [4.4.26]$$

The Jacobian of this matrix is

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{G}}{\partial \theta_1} & \frac{\partial \mathbf{G}}{\partial \theta_2} & \frac{\partial \mathbf{G}}{\partial \theta_3} \end{bmatrix}. \qquad [4.4.27]$$

Performing the differentiation yields

$$J = \begin{bmatrix} -\frac{c}{m^2} & \frac{1}{m} & 0 \\ -\frac{k}{m^2}+\frac{2c^2}{m^3} & -\frac{2c}{m^2} & \frac{1}{m} \\ \frac{4kc}{m^3}-\frac{3c^3}{m^4} & -\frac{2k}{m^2}+\frac{3c^2}{m^3} & -\frac{2c}{m^2} \\ \frac{2k^2}{m^3}-\frac{9c^2k}{m^4}+\frac{4c^4}{m^5} & \frac{6ck}{m^3}-\frac{4c^3}{m^4} & -\frac{2k}{m^2}+\frac{3c^2}{m^3} \end{bmatrix} \qquad [4.4.28]$$

This entire matrix may be multiplied by m to more easily show the identifiability problem. After the multiplication, the parameters m, c, and k appear only in the combinations $\left(\frac{c}{m}\right)$ and $\left(\frac{k}{m}\right)$. The new matrix is

$$Jm = \begin{bmatrix} -\left(\frac{c}{m}\right) & 1 & 0 \\ -\left(\frac{k}{m}\right)+2\left(\frac{c}{m}\right)^2 & -2\left(\frac{c}{m}\right) & 1 \\ 4\left(\frac{k}{m}\right)\left(\frac{c}{m}\right)-3\left(\frac{c}{m}\right)^3 & -2\left(\frac{k}{m}\right)+3\left(\frac{c}{m}\right)^2 & -2\left(\frac{c}{m}\right) \\ 2\left(\frac{k}{m}\right)^2-9\left(\frac{c}{m}\right)^2\left(\frac{k}{m}\right)+4\left(\frac{c}{m}\right)^4 & 6\left(\frac{c}{m}\right)\left(\frac{k}{m}\right)-4\left(\frac{c}{m}\right)^3 & -2\left(\frac{k}{m}\right)+3\left(\frac{c}{m}\right)^2 \end{bmatrix} \qquad [4.4.29]$$

The two ratios of three parameters appearing in this matrix suggest that the only two parameters can be identified. Unfortunately, it can be difficult to show J has rank of two. Three general methods were examined to compute the rank of the matrix.

First, the determinant of J could be examined algebraically using algebraic manipulators such as MACSYMA or SMP. The algebra would eventually show the

determinant to be zero for all nonzero values of the parameters. However, as the complexity of the model increased, the algebraic manipulation would become prohibitively complex. Algebraic manipulators also require a large amount of time to become proficient using them. Because of these limitations, the algebraic manipulation approach was not used in this work.

Next, the determinant of the matrix may be calculated directly. If the determinant of the matrix is zero, then the matrix has less than full rank. Unfortunately, numerical difficulties often arise when calculating the determinant. The resolution of the computer limits the quality of the rank determination. To test if the resolution of the computer was affecting the result, the determinant of **J** was performed in single, double, and quad precision. The magnitudes of the determinant from this test are shown in Table 6. This table shows the computer resolution limited the value returned for the determinant. Because of these numerical difficulties met, this method was not used.

Finally, a commercial software package Matlab (Matlab Users Guide, 1987) was used to determine the rank of the Jacobian matrix. Matlab contains a function to compute the rank of a matrix. Matlab computes the rank by determining the number of singular values of **J** that are larger than δ in the equation

$$\delta = \max\{\text{size}(\mathbf{J})\}\text{norm}(\mathbf{J})\varepsilon \qquad [4.4.30]$$

where

$\max\{\text{size}(\mathbf{J})\}$	=	largest dimension of matrix **J**
norm(**J**)	=	largest singular value of matrix **J**
ε	=	precision of specific machine used

This procedure was very stable and used information from the machine precision. Thus, Matlab was used to determine the rank of the matrix.

Table 6. Effect of computer precision on determinant

Precision	*Magnitude of Determinant*
single	10^{-5}
double	10^{-15}
quad	10^{-30}

Using Matlab, the rank of the displacement-input system was determined to be two for almost all combinations of parameter values. The only exceptions occurred when parameters of widely different magnitudes (about 10^8) were used. The matrix was rank one here. These exceptions would not occur in any physically significant model.

The implication of the rank equal to two is that only two of the three parameters may be identified. In the estimation, seen earlier in section [4.3.3], the three parameter values returned from the estimation were linearly dependent. Therefore, only two parameters such as damping ratio and natural frequency were determined.

Thus, Grewal and Glover's method may be used to determine the identifiability of state-space dynamic models. However, even for the one-degree-of-freedom model

presented here, the method can become quite complex. The use of commercial computer programs can reduce this difficulty.

4.4.2.2 Reid's Method

Reid's method of identifiability for the displacement-input case is very similar to the formulation of the force-input case. For both cases, the **A** matrix is equal and therefore contains the same eigenvalue and eigenvector sensitivities. The differences arise in the formulation of the **B** and **C** matrices. For the displacement-input case, the partial derivatives of the matrices **B** and **C** are

$$\frac{\partial \mathbf{B}}{\partial \theta_1} = \frac{\partial}{\partial k}\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \qquad [4.4.31]$$

$$\frac{\partial \mathbf{B}}{\partial \theta_2} = \frac{\partial}{\partial c}\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \qquad [4.4.32]$$

$$\frac{\partial \mathbf{B}}{\partial \theta_3} = \frac{\partial}{\partial m}\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \qquad [4.4.33]$$

$$\frac{\partial \mathbf{C}}{\partial \theta_1} = \frac{\partial}{\partial k}\begin{bmatrix}\frac{k}{m} & \frac{c}{m}\end{bmatrix} = \begin{bmatrix}\frac{1}{m} & 0\end{bmatrix} \qquad [4.4.34]$$

$$\frac{\partial \mathbf{C}}{\partial \theta_2} = \frac{\partial}{\partial c}\begin{bmatrix}\frac{k}{m} & \frac{c}{m}\end{bmatrix} = \begin{bmatrix}0 & \frac{1}{m}\end{bmatrix} \qquad [4.4.35]$$

$$\frac{\partial \mathbf{C}}{\partial \theta_3} = \frac{\partial}{\partial m}\begin{bmatrix}\frac{k}{m} & \frac{c}{m}\end{bmatrix} = \begin{bmatrix}-\frac{k}{m^2} & -\frac{c}{m^2}\end{bmatrix} \qquad [4.4.36]$$

The sensitivity matrix $\mathbf{G}_{zs}$ was then assembled from equation [4.4.11]. Once the matrix was assembled, its rank was determined using Matlab. Similar numerical difficulties were encountered with Reid's method as with Grewal and Glover's method in finding the rank of the sensitivity matrix. In order to make the determination of the rank a more stable procedure, a matrix identity from Chen (1970) was used. This matrix identity states that the rank of a rectangular nxm matrix is equal to the rank of that matrix times its transpose. Therefore, the equation used was

$$\text{rank}(\mathbf{G}_{zs}) = \text{rank}(\mathbf{G}_{zs}\mathbf{G}'_{zs}) \qquad [4.4.37]$$

This equation better conditioned the numerical procedure. Before the relation from Chen was used, Matlab would occasionally compute a value of three for the rank. After this change, the rank of $\mathbf{G}_{zs}$ was determined to be two for all combinations of values checked in the region near the initial values of the estimation. Therefore, Reid's method indicates the displacement-input model is identifiable for only two of the three system parameters.

4.5 *Conclusions*

This chapter contains parameter estimation examples for two single degree-of-freedom spring-mass-damper systems. Only one of the systems is identifiable. The test problem illustrated the pitfalls of estimating nonidentifiable systems. The difficulties encountered in estimating these basic systems provides considerable insight into the identification of more complex dynamic systems. For example, the convergence of the performance index cannot be fully relied upon to

indicate that the parameters have each converged to their final values. Also, the Bayesian term can occlude estimation results of nonidentifiable systems. Because the rail vehicle model estimated in this thesis closely resembles combinations of the test problems, the conclusions from the identifiability of the test problems are directly applicable to the rail vehicle estimation.

This chapter also contains evaluations of two distinct methods to determine the identifiability of linear dynamic systems. Each method has several advantages and disadvantages, but both provided consistent results about the identifiability of the systems tested. These methods both provide identifiability information from system models, and if they are applied prior to experimentation, they can decrease the likelihood of fundamental estimation errors.

Finally, this chapter contains evaluations of the estimation and data processing techniques. All of the rail vehicle techniques will closely follow the methods detailed in this chapter.

Chapter 5

Experimental Setup and Data Processing

5.1 *Transportation Test Center*

The rail vehicle in this thesis was tested in the Rail Dynamics Laboratory (RDL) at the Transportation Test Center (TTC) in Pueblo, Colorado during August and September of 1985. The TTC provides "practical research and development testing of railroad systems, transit systems, and other ground transportation concepts with the objective of promoting a safe, adequate, economical, and efficient national transportation system" (RDL Users Guide, 1978). The Vibration Test Unit (VTU), a major testing component of the RDL, provides vertical and lateral excitation to the vehicle wheels (Inskeep and Roberts, 1982). All of the test data used in this thesis was performed on the VTU.

5.2 *Vibration Test Unit*

Inskeep and Roberts (1982) described many of the capabilities of the VTU. The VTU can subject a stationary rail vehicle to controlled vertical and lateral inputs at the wheels, simulating the dynamic effects of perturbed track. Hydraulic shakers on the VTU apply the rail inputs. Bearing assemblies connect the car wheels to the shakers and allow simultaneous vertical and lateral input motions and wheelset rotations.

A computer control system provides displacement time histories to the actuators. These time histories may be sinusoidal, impulse, or random forcing functions in virtually any combination of vertical and lateral motions. Figure 23 shows the frequency and displacement performance of the actuators. Beyond the upper test frequency, waveform distortions occur and phase correlations between the shakers rapidly deteriorate.

5.3 *Experimental Testing*

The TTC investigated eight modes of vibration for the rail vehicle both with and without a road trailer. The five rigid body modes were bounce, pitch, roll, sway, and yaw. The three flexible body modes were vertical bending, lateral bending, and twist. The flexible body modes occur at much higher frequencies than the rigid body modes. Negligible coupling exists between the rigid and the flexible body modes. For this thesis, only the unloaded rail vehicle tests were considered. Also, the only

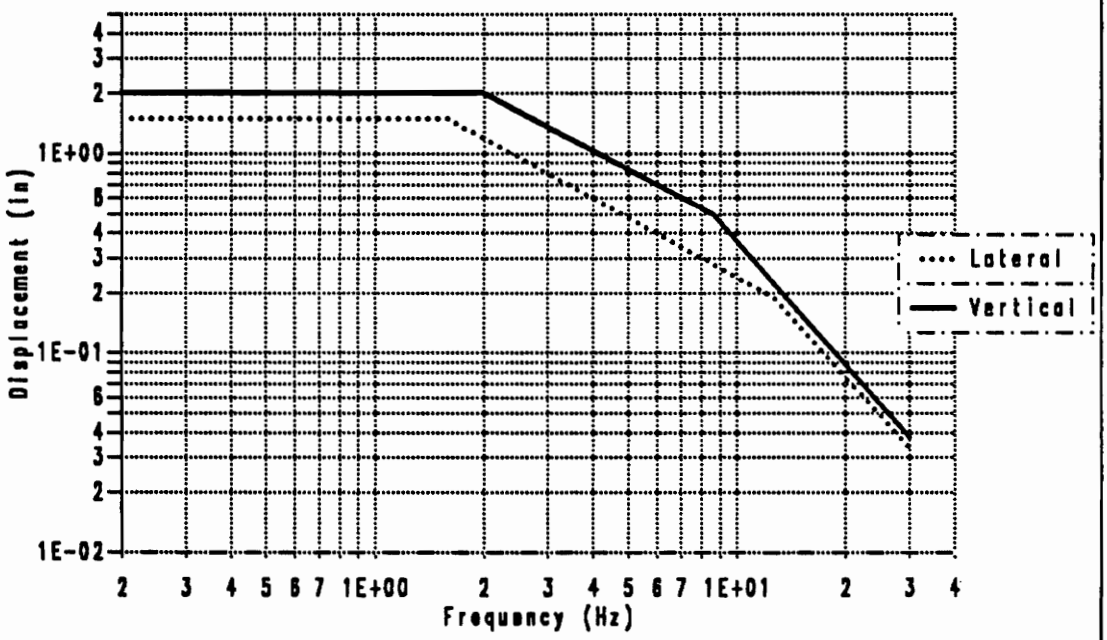


Figure 23. Actuator characteristics

rigid body modes investigated were bounce, pitch, and roll. Figure 24 and Table 7 describe the actuator magnitude and phase relationships for these modes.

The vehicle was mounted on the VTU to provide similar operating conditions with the actual environment. The vehicle wheels rested on a short section of rail mounted to each actuator. A coupler between the vehicle and ground restrained the vehicle from large lateral deflections. This experimental setup allowed realistic force transmission to excite the rail vehicle during the testing.

5.4 Data Acquisition

5.4.1 Instrumentation

The data acquisition system recorded and processed 51 channels of information for the tests with the road trailer attached and 35 channels for the rail car alone. The dynamic characteristics of the actuators did not attenuate the results because the testing occurred at low frequency. The data was sampled at 150 Hz, more than 10 times faster than the highest input frequency.

Four different types of transducers measured the response of the rail vehicle. The transducers used were

1. string-pot displacement transducers
2. accelerometers
3. rotational gyrometers
4. strain gauges

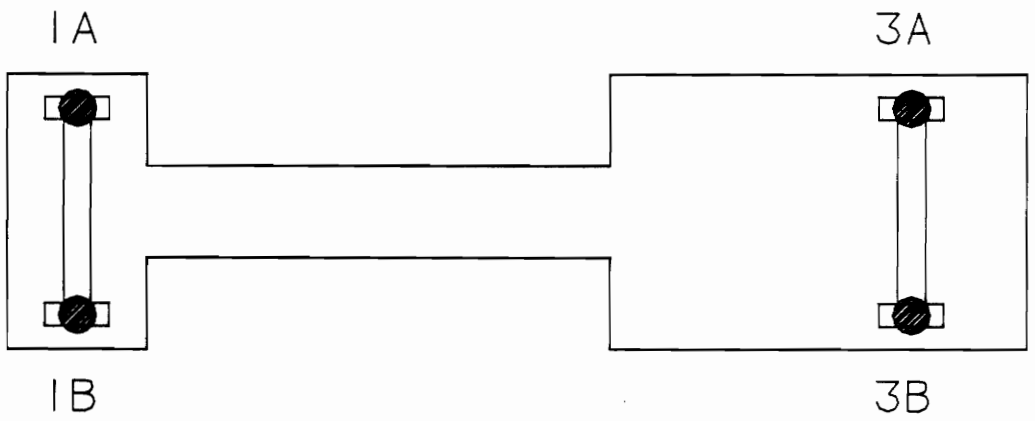


Figure 24. Actuator locations

Table 7. Actuator relative amplitude and phase

Test	*Actuator Relative Amplitude and Phase (-1 = 180 ° out of phase*				*Frequency Range (Hz)*
	1A	*1B*	*3A*	*3B*	
Bounce	1	1	1	1	1-5
Pitch	1	1	-1	-1	1-5
Roll	1	-1	1	-1	3.5-8

These transducers measured linear displacement, acceleration, rotational displacement, and force, respectively. The strain gauges were mounted on the rails to measure the input force to the rail vehicle wheels.

5.4.2 Locations of Instrumentation

Table 8 lists the numbers, names, and engineering units for each channel recorded during the testing. The inactive channels all measure the road trailer response when it is mounted on the rail vehicle. Figure 25 shows only the locations of transducers used for data processing in this thesis.

Table 8. Vibration Test Unit: Instrumentation Labels

Channel Number	Channel Name	Channel Units	Active	Channel Number	Channel Name	Channel Units	Active
1	1AZ	in	Yes	27	A27Z	g	Yes
2	1BZ	in	Yes	28	A28Z	g	Yes
3	1CX	in	Yes	29	A29X	g	Yes
4	3AZ	in	Yes	30	A30Z	g	Yes
5	3BZ	in	Yes	31	A31Z	g	Yes
6	3CX	in	Yes	32	A32X	g	Yes
7	RF1AZ	kip	Yes	33	A33Z	g	Yes
8	RF1AX	kip	Yes	34	A34Z	g	Yes
9	RF1BZ	kip	Yes	35	A35X	g	Yes
10	RF1BX	kip	Yes	36	1HND	UE	No
11	RF3AZ	kip	Yes	37	A37Z	g	No
12	RF3AX	kip	Yes	38	A38X	g	No
13	RF3BZ	kip	Yes	39	A39Z	g	No
14	RF3BX	kip	Yes	40	A40Z	g	No
15	D15Z	in	Yes	41	A41X	g	No
16	D16Z	in	Yes	42	A42Z	g	No
17	D17X	in	Yes	43	A43X	g	No
18	D18Z	in	Yes	44	GR1	deg	Yes
19	D19Z	in	Yes	45	GR2	deg	Yes
20	D20X	in	Yes	46	GR3	deg	No
21	D21Z	in	No	47	S47X	193.2	Yes
22	D22Z	in	No	48	S48X	193.2	Yes
23	D23Z	in	No	49	S49Y	193.2	Yes
24	D24Z	in	No	50	D50Z	in	No
25	D25X	in	No	51	MARKE	V	No
26	D26X	in	No				

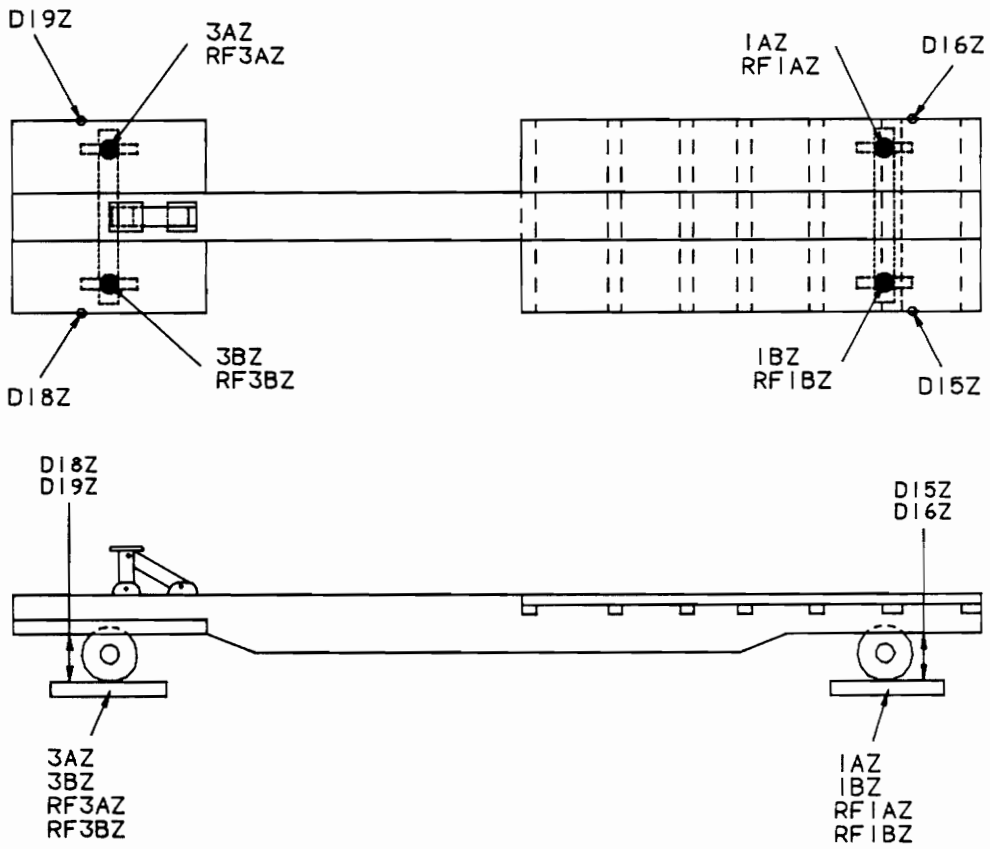


Figure 25. Transducer locations on rail vehicle

5.4.3 Preliminary Data Processing and Recording

Engineers at the RDL used a PDP-11/60 computer to perform all of the data acquisition and preliminary processing tasks. The analog values were converted to a binary equivalent, multiplied by the appropriate scaling factor, and stored in the computer memory for further processing. The TTC sent the data for this thesis on tapes from the originally stored data. Thus, the data was received in engineering units ready for processing by methods similar to those used in Chapter 4.

5.5 Data Processing

5.5.1 Initial Data Conditioning

The first processing of the results from the TTC was identical to the processing used in the test case. The steps were

1. Transform data to mean of zero by removing offset and trend.
2. Discard data further than 4 standard deviations from the mean.

This processing removed any drift from the instrumentation and deleted data affected by random disturbances during the experiment.

5.5.2 Determining Frequency Content of Input Signal

Before further processing the data, the frequency content of the input signal was determined. The frequency content of the signal provided insight into the characteristics of the frequency response function. For example, if the testing frequencies missed the resonance peaks completely, then the FRFs would be inaccurate in that region.

A zero-crossing method determined the frequency content of the input signal. This method calculated the change in time between zero amplitude crossings of the input signal and averaged two crossings for a nominal cycle frequency. The equation for the cycle frequency is

$$f_i = \frac{1}{\frac{1}{2}\left[(T_i - T_{i-1}) + (T_{i+1} - T_i)\right]} \qquad [5.1.1]$$

or

$$f_i = \frac{2}{T_{i+1} - T_{i-1}} \qquad [5.1.2]$$

where

f_i = frequency of the i^{th} cycle

T_i = zero crossing time at middle of i^{th} cycle

T_{i+1} = zero crossing time at end of i^{th} cycle

T_{i-1} = zero crossing time at beginning of i^{th} cycle

The T_is were calculated using a linear interpolation method between the two data points on either side of the zero crossing.

The frequency content of the bounce test input signal is shown in Figure 26. The data originally resembled a sine-sweep through the test frequency range. However, expanding the figure around 3 Hz shows a ramp-dwell sweep. Figure 27 shows how the frequency ramps up from plateau to plateau. The frequency increases approximately 0.125 Hz between each plateau. This analysis shows there are not equal amounts of information at all frequencies. Therefore, the computed FRFs may contain poorer estimates at the frequencies ramped through than at the dwell locations.

5.5.3 Computing Frequency Response Functions

Several different FRF computation methods were examined to determine the effects of input frequency, noise, and data record length. The methods examined were

1. Discrete Fourier transform ratio method
2. Fast Fourier transform ensemble average method

These methods were then compared to the in-house computation method used at the Transportation Test Center.

5.5.3.1 Discrete Fourier Transform Method

Because the frequency content of the input signal contains many dwell locations, the data can be broken down into discrete sections with one input-output section containing a single frequency. This analysis of the data implies the use of

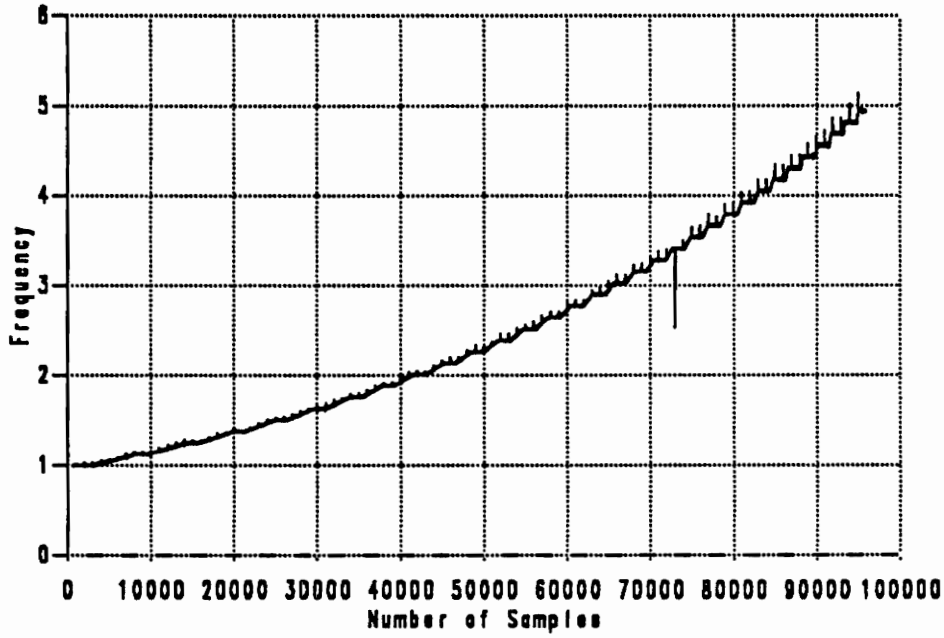


Figure 26. Frequency content of the input signal

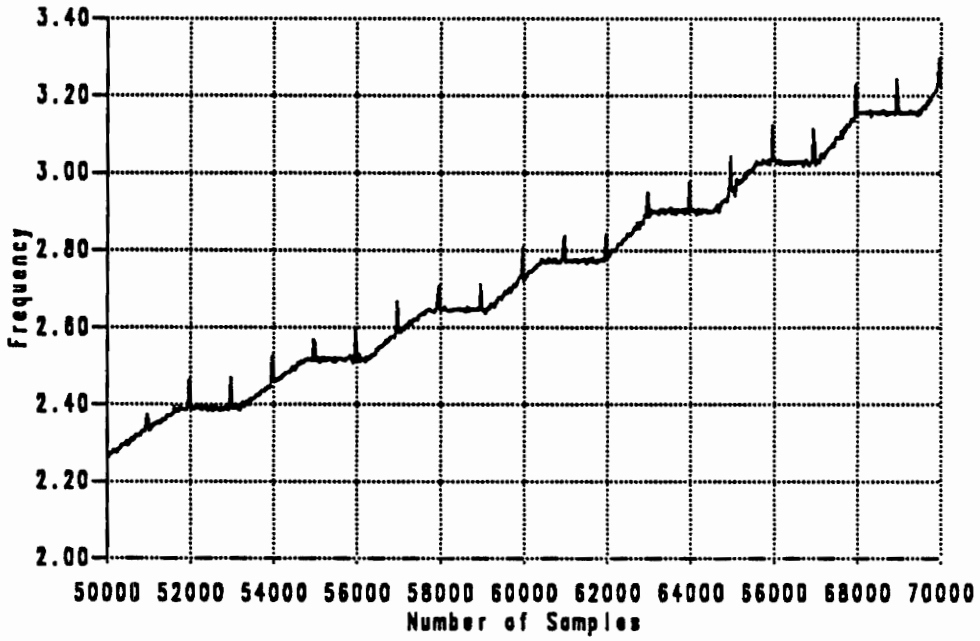


Figure 27. Expanded frequency content of the input signal near resonance

Discrete Fourier Transforms (DFTs) to determine the FRF because the DFT computed in each dwell could contain all of the information in one frequency bin.

The DFT algorithm must be applied to all dwell locations to obtain each FRF. The algorithm may be summarized in the following steps:

1. Determine location of each dwell from the input frequency signal.
2. Compute DFT in each dwell location of both input and output signals.
3. Compute input-output ratio for the dwell frequency bin of each DFT.

This algorithm therefore requires computation of DFTs for both the input and the output signal in each frequency dwell. This process requires much computing to determine one FRF.

Figure 28 (a) and (b) shows the DFT of the input and output signal for one frequency dwell near the resonant frequency of the bounce test. Each frequency bin represents 0.125 Hz. Because the input and output signals contain only one frequency, the useful information in these graphs lies in bin 23.

The ratio of the output to input frequency magnitude is computed for the bin containing the dominant frequency. Here, the calculation is

$$H_i = \frac{Mo_i}{Mi_i} = \frac{2.96}{0.734} = 4.03 \qquad [5.1.3]$$

where

H_i = FRF magnitude at i^{th} frequency

Mo_i = magnitude of output DFT at i^{th} frequency

Mi_i = magnitude of input DFT at i^{th} frequency

Thus, from Figure 28 one value from frequency bin 23 (2.88 Hz) is computed for the FRF. After this step is completed, step 2 and 3 above are repeated throughout the frequency range. Figure 29 shows an entire FRF computed using the DFT method.

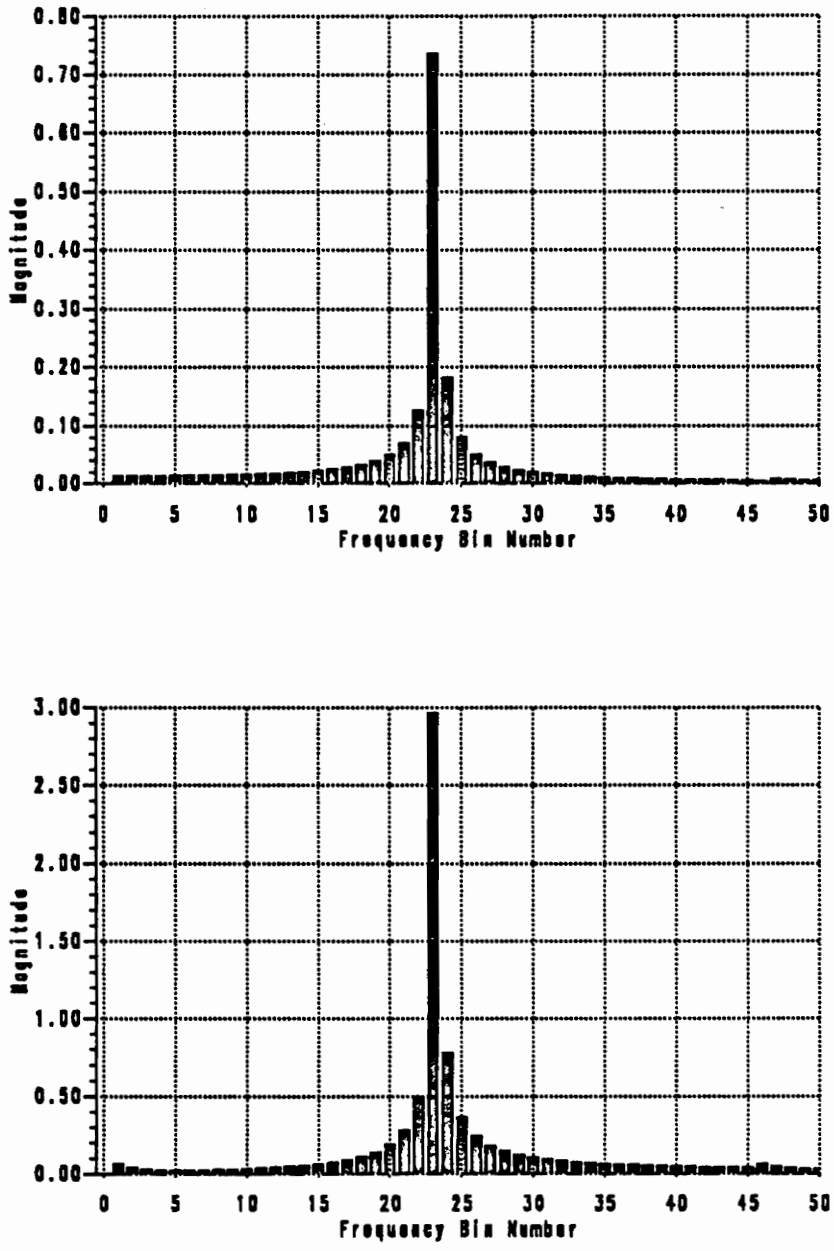


Figure 28. DFT of (a) input and (b) output signals for bounce test

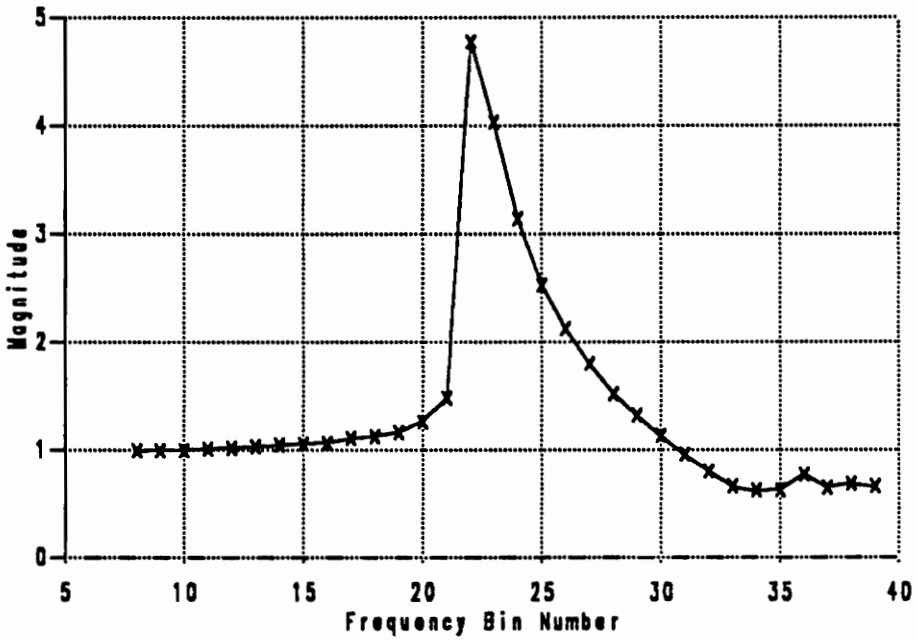


Figure 29. Frequency response function computed using DFT method

The DFT method contains several advantages and disadvantages. One advantage is the method reduces the huge quantity of time response data to only 20 - 30 pieces. Also, the computation is only performed in a frequency dwell thus removing transients from the data. Finally, the DFT response is independent of record length. Unfortunately, the DFT method requires large amounts of computer processing time. These large computing requirements result from the computation of many DFTs for one FRF. Another disadvantage is the frequency content of the signal does not always fall into only one bin of the DFT and can spill into several adjacent bins. This problem complicates the simple ratio method to determine the FRF magnitude at one frequency. Finally, the DFT method cannot compensate for measurement or process noise in the system. These disadvantages limit the usefulness of the DFT method.

5.5.3.2 Fast Fourier Transform Ensemble Average Method

The FFT method of ensemble averaging the data records was also investigated using the same algorithms developed in Chapter 4. Implementation of this method is much easier than the DFT method. This method also requires much less processing time. However, the results differ significantly depending on both record length and FRF computation method ($\mathbf{H}_1$, $\mathbf{H}_2$, or $\mathbf{H}_3$).

Three common record lengths used with the FFT method are 1024, 2048, and 4096. Figures 30-32 show the FRF differences for different record lengths. These figures show a wide variance in FRF depending on record length used. As the record length of the data decreases, the frequency resolution of the FRF decreases. The decreased resolution raises the probability of the FRF missing the resonant frequencies, especially with a lightly damped system. From this analysis, the FRF

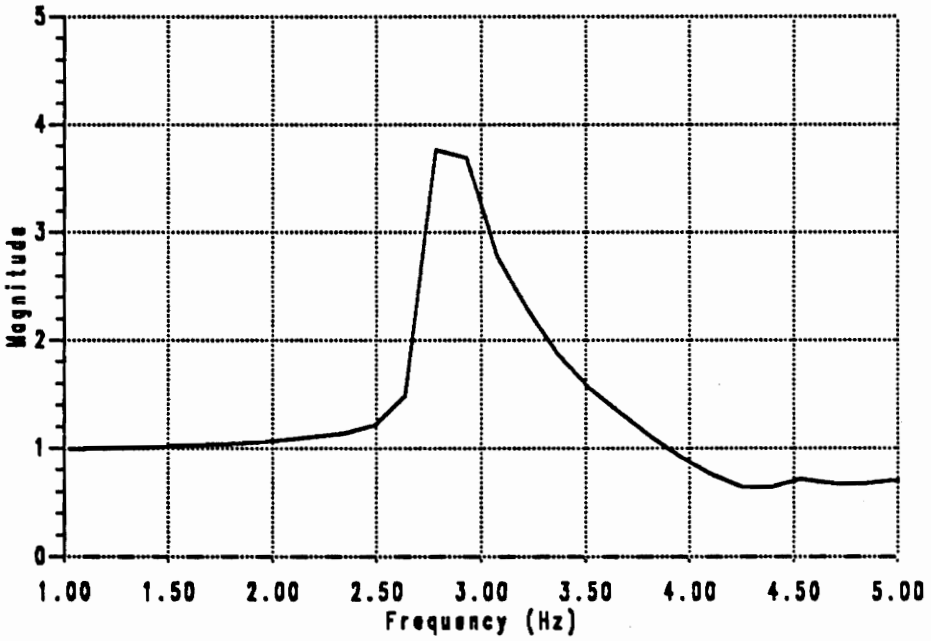


Figure 30. FRF using FFT method with record length of 1024

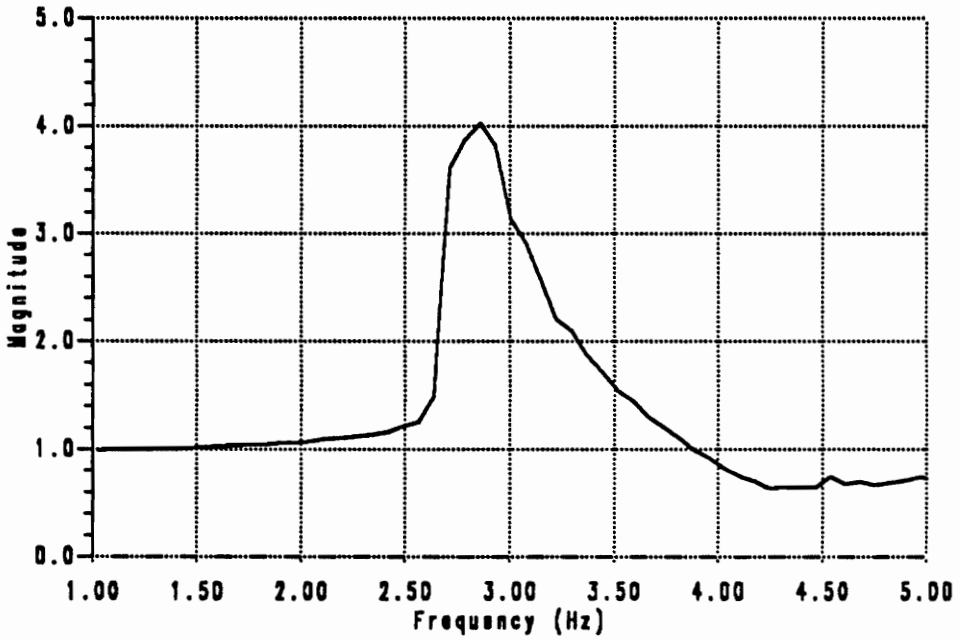


Figure 31. FRF using FFT method with record length of 2048

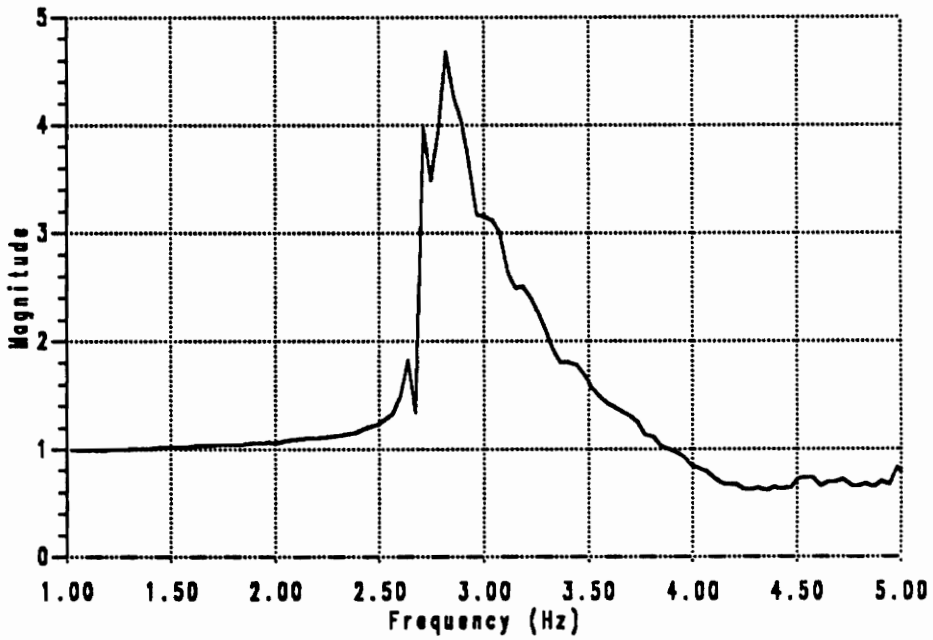


Figure 32. FRF using FFT method with record length of 4096

using a record length of 4096 provides a higher estimate of the FRF and is consistent with the results from the DFT method. Therefore, a record length of 4096 was used in the computation of the FRFs using the FFT method.

When using the FFT method, the choice of computation scheme, H_1 , H_2, or H_3, must still be considered. Chapter 4 discussed the differences between the two methods. Figure 33 shows the differences between the H_1 and H_2 methods. H_3 is simply the average of H_1 and H_2. Both of the methods provided similar results because of a clean, well conditioned input signal. However, H_1 is less sensitive to process noise than H_2 Therefore, since the measurement noise is very low and H_1 is less sensitive to process noise than H_2 , the H_1 method was chosen.

The advantages of the FFT ensemble averaging method are twofold. First, the method is easy to implement and is not highly computing intensive. Second, the FFT method contains several different variations to compensate for peak width, measurement noise, and process noise. One disadvantage, however, is the sensitivity of the FFT method to the record length used.

5.5.3.3 Transportation Test Center Computation Method

Finally, the Transportation Test Center also computed FRFs of the vehicle response data. The engineers at the TTC used an in-house computer program to compute both the magnitude and phase of the FRFs. Unfortunately, the exact method they used to create the FRFs is unknown. However, their graphs show good agreement between their computation methods and the methods described above. Figure 34 shows the TTC computation of the same FRF computed above. The TTC data shows a higher peak value at the natural frequency but with that exception, the overall response is similar.

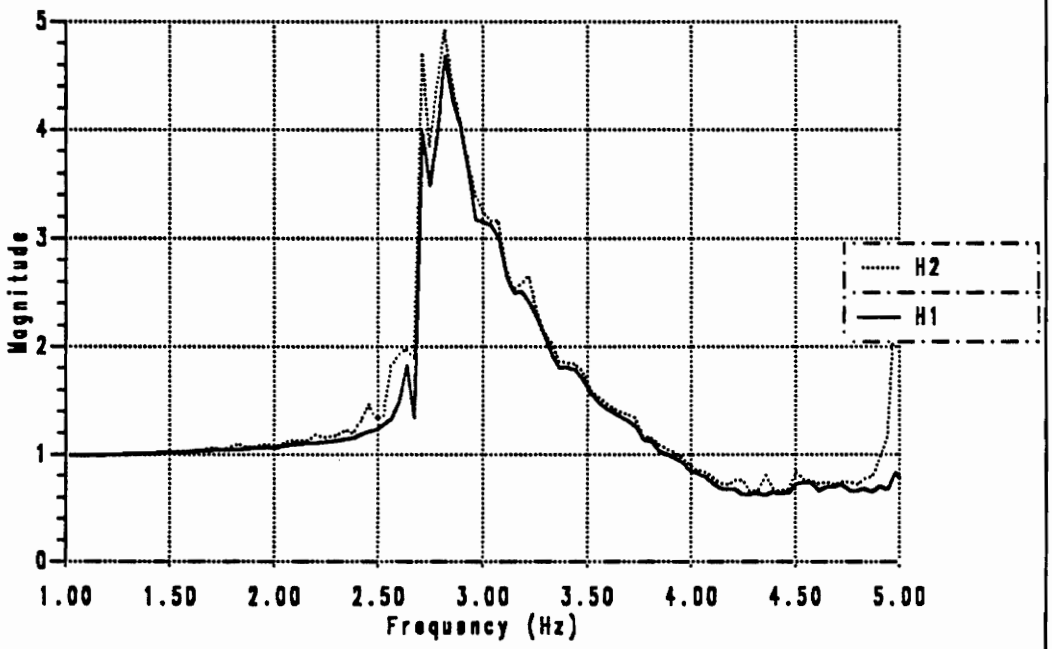


Figure 33. Comparison between H_1 and H_2

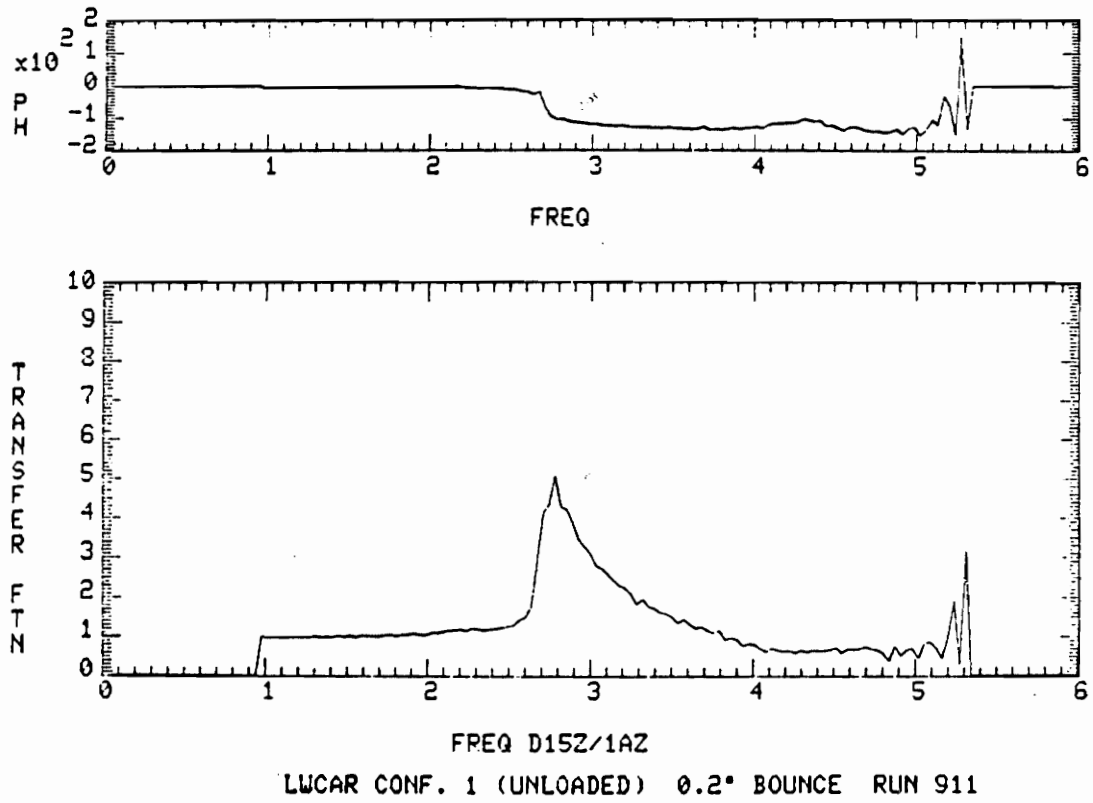


Figure 34. Transportation Test Center frequency response function computation

5.5.4 Choice of Computation Method

Table 9 shows the comparison between the DFT and FFT methods. The FFT method is more advantageous than the DFT method because of easy implementation and accommodation of noisy signals. Thus, the FFT method was used to process all of the rail vehicle data.

Table 9. Comparison of FRF computation methods

Method	*Advantages*	*Disadvantages*
DFT	FRF independent of record length reduces to small amount of data no transients in data	difficult to implement large computing requirements no compensation for noise
FFT	easy to implement low computing requirements noise compensation	FRF highly sensitive to record length

Chapter 6

Model Description and Estimation

6.1 *Introduction*

The previous chapters provided theoretical background and development of parameter estimation methods, the identifiability problem, and data processing techniques. These three areas are now combined to estimate the parameters for the rail vehicle system.

6.2 *General Description of the Vertical Vehicle Model*

The vehicle investigated was a lightweight single axle car with a European International Union of Railways (UIC) suspension system. Only the vertical dynamics

of the rail vehicle were considered. Thus, data from the bounce, pitch and roll tests provided all of the results.

Wormley and Tombers (1983) described the suspension system. They stated "the single axle is suspended via the bearing box with a leaf vertical suspension to carry the vertical loads." Therefore in the vertical direction the load is carried only through a primary suspension, not a primary and secondary suspension combination typical of most railcars. Figure 35 shows the single axle truck configuration (Irani, et al., 1986).

The main structural element of the vehicle is the lightweight center sill. The sill resembles a hollow beam and is flexible in torsion. When attached, the road trailer snaps into the kingpin connection on the railcar and its wheels rest on the plates to the side of the rail vehicle.

6.3 *Building the Model*

6.3.1 Assumptions

The vertical dynamics model incorporated several assumptions. The assumptions were:

1. The suspension contains only linear elements.
2. There is no flexible body motion.
3. The vehicle and suspension are symmetric.

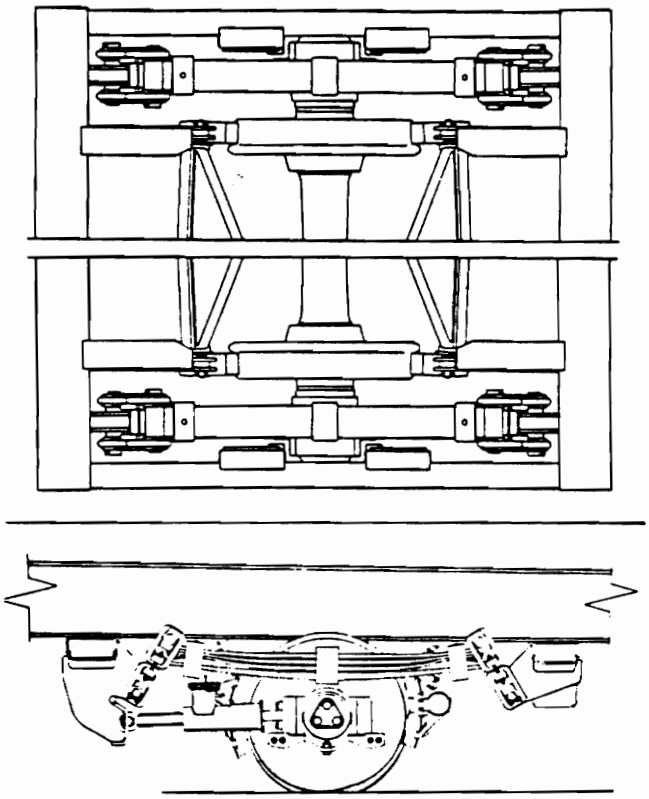


Figure 35. Single axle truck configuration

Though these assumptions are always inaccurate to some degree, they are adequate for a first estimation attempt.

6.3.2 Derivation of Vertical Equations of Motion

The previous assumptions simplify the derivation of the vertical vehicle model. Figure 36 shows one side of the vehicle. From Figure 36, the vehicle equations of motion can be derived using Newton's laws. By summing the forces and moments for each mode, the equations in matrix form are:

$$[\mathbf{M}]\begin{bmatrix}\ddot{Y}\\ \ddot{\Psi}\\ \ddot{\Theta}\end{bmatrix} + [\mathbf{C}]\begin{bmatrix}\dot{Y}\\ \dot{\Psi}\\ \dot{\Theta}\end{bmatrix} + [\mathbf{K}]\begin{bmatrix}Y\\ \Psi\\ \Theta\end{bmatrix} = [\mathbf{B_1}]\begin{bmatrix}\dot{z}\\ \dot{\alpha}\\ \dot{\phi}\end{bmatrix} + [\mathbf{B_2}]\begin{bmatrix}z\\ \alpha\\ \phi\end{bmatrix} \qquad [6.3.1]$$

where

Y	=	vertical output
Ψ	=	pitch angle output
Θ	=	roll angle output
z	=	vertical input
α	=	pitch angle input
ϕ	=	roll angle input

The matrices for the general system without suspension symmetry are

$$[\mathbf{M}] = \begin{bmatrix} m & 0 & 0 \\ 0 & J_x & 0 \\ 0 & 0 & J_z \end{bmatrix} \qquad [6.3.2]$$

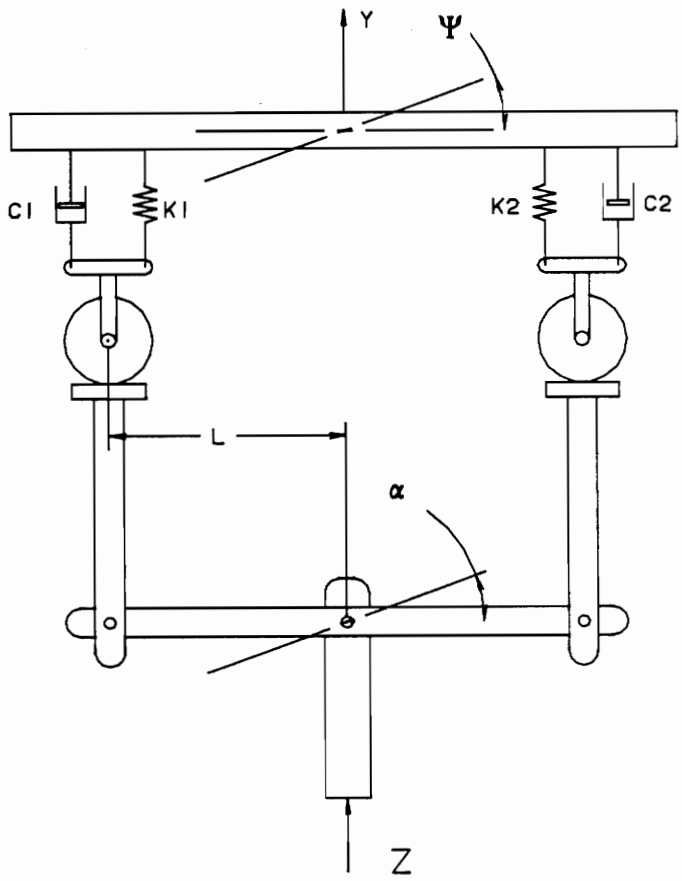


Figure 36. Simplified vehicle structure (side view)

$$[C] = \begin{bmatrix} (c_1 + c_2 + c_3 + c_4) & (c_1 + c_2 - c_3 - c_4)\,l & (-c_1 + c_2 - c_3 + c_4)\,d_2 \\ (c_1 + c_2 - c_3 - c_4)\,l & (c_1 + c_2 + c_3 + c_4)\,l^2 & (-c_1 + c_2 + c_3 - c_4)\,l\,d_2 \\ (-c_1 + c_2 - c_3 + c_4)\,d_2 & (-c_1 + c_2 + c_3 - c_4)\,l\,d_2 & (c_1 + c_2 + c_3 + c_4)\,d_2^2 \end{bmatrix} \qquad [6.3.3]$$

$$[K] = \begin{bmatrix} (k_1 + k_2 + k_3 + k_4) & (k_1 + k_2 - k_3 - k_4)\,l & (-k_1 + k_2 - k_3 + k_4)\,d_2 \\ (k_1 + k_2 - k_3 - k_4)\,l & (k_1 + k_2 + k_3 + k_4)\,l^2 & (-k_1 + k_2 + k_3 - k_4)\,l\,d_2 \\ (-k_1 + k_2 - k_3 + k_4)\,d_2 & (-k_1 + k_2 + k_3 - k_4)\,l\,d_2 & (k_1 + k_2 + k_3 + k_4)\,d_2^2 \end{bmatrix} \qquad [6.3.4]$$

$$[B_1] = \begin{bmatrix} (c_1 + c_2 + c_3 + c_4) & (c_1 + c_2 - c_3 - c_4)\,l & (-c_1 + c_2 - c_3 + c_4)\,d_2 \\ (c_1 + c_2 - c_3 - c_4)\,l & (c_1 + c_2 + c_3 + c_4)\,l^2 & (-c_1 + c_2 + c_3 - c_4)\,l\,d_2 \\ (-c_1 + c_2 - c_3 + c_4)\,d_2 & (-c_1 + c_2 + c_3 - c_4)\,l\,d_2 & (c_1 + c_2 + c_3 + c_4)\,d_2^2 \end{bmatrix} \qquad [6.3.5]$$

$$[B_2] = \begin{bmatrix} (k_1 + k_2 + k_3 + k_4) & (k_1 + k_2 - k_3 - k_4)\,l & (-k_1 + k_2 - k_3 + k_4)\,d_2 \\ (k_1 + k_2 - k_3 - k_4)\,l & (k_1 + k_2 + k_3 + k_4)\,l^2 & (-k_1 + k_2 + k_3 - k_4)\,l\,d_2 \\ (-k_1 + k_2 - k_3 + k_4)\,d_2 & (-k_1 + k_2 + k_3 - k_4)\,l\,d_2 & (k_1 + k_2 + k_3 + k_4)\,d_2^2 \end{bmatrix} \qquad [6.3.6]$$

Note $[B_1] = [C]$ and $[B_2] = [K]$. Implementing the symmetry assumption, the damping and stiffness become

$$k = k_1 = k_2 = k_3 = k_4 \qquad [6.3.7]$$

$$c = c_1 = c_2 = c_3 = c_4 \qquad [6.3.8]$$

This reduces the matrices to

$$[C] = \begin{bmatrix} 4c & 0 & 0 \\ 0 & 4c\,l^2 & 0 \\ 0 & 0 & 4c\,d_2^2 \end{bmatrix} \qquad [6.3.9]$$

$$[K] = \begin{bmatrix} 4k & 0 & 0 \\ 0 & 4k\,l^2 & 0 \\ 0 & 0 & 4k\,d_2^2 \end{bmatrix} \qquad [6.3.10]$$

$$[B_1] = \begin{bmatrix} 4c & 0 & 0 \\ 0 & 4cl^2 & 0 \\ 0 & 0 & 4cd_2^2 \end{bmatrix} \qquad [6.3.11]$$

$$[B_2] = \begin{bmatrix} 4k & 0 & 0 \\ 0 & 4kl^2 & 0 \\ 0 & 0 & 4kd_2^2 \end{bmatrix} \qquad [6.3.12]$$

Because of the assumed vehicle symmetry, the bounce, pitch, and roll modes uncouple. Therefore, the vehicle model decomposes into three single degree-of-freedom models.

6.3.2.1 Equations of Motion for the Bounce Mode

Figure 37 shows the single degree-of-freedom decomposition of the vehicle model for the bounce mode. The equation of motion for the bounce mode is

$$m\ddot{Y} + (4c)\dot{Y} + (4k)Y = (4c)\dot{z} + (4k)z \qquad [6.3.13]$$

where

m = vehicle mass

c = linear damping constant for each suspension group

k = linear spring constant for each suspension group

This model is equivalent to the displacement-input model from Chapter 4. The state space implementation of this equation is

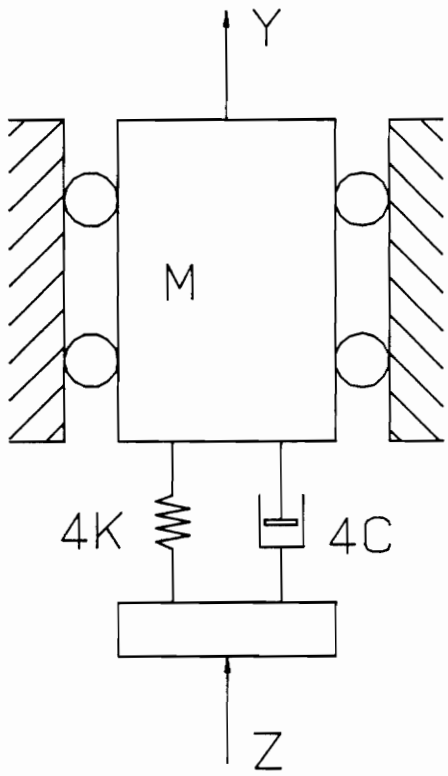


Figure 37. Single degree-of-freedom model of the bounce mode

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\left(\frac{4k}{m}\right) & -\left(\frac{4c}{m}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} z \qquad [6.3.14]$$

$$y = \begin{bmatrix} \left(\frac{4k}{m}\right) & \left(\frac{4c}{m}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \qquad [6.3.15]$$

6.3.2.2 Equations of Motion for the Pitch and Roll Modes

Figure 38 shows a representation of the rotational modes. The equation of motion for the pitch mode is

$$J_x\ddot{\Psi} + (4cl^2)\dot{\Psi} + (4kl^2)\Psi = (4cl^2)\dot{\alpha} + (4kl^2)\alpha \qquad [6.3.16]$$

where

J_x = vehicle pitch moment of inertia.

The equation of motion for the roll mode is

$$J_z\ddot{\Theta} + (4cd_2^2)\dot{\Theta} + (4kd_2^2)\Theta = (4cd_2^2)\dot{\phi} + (4kd_2^2)\phi \qquad [6.3.17]$$

where

J_z = vehicle roll moment of inertia.

The state-space formulation of these modes is similar to the bounce mode implementation.

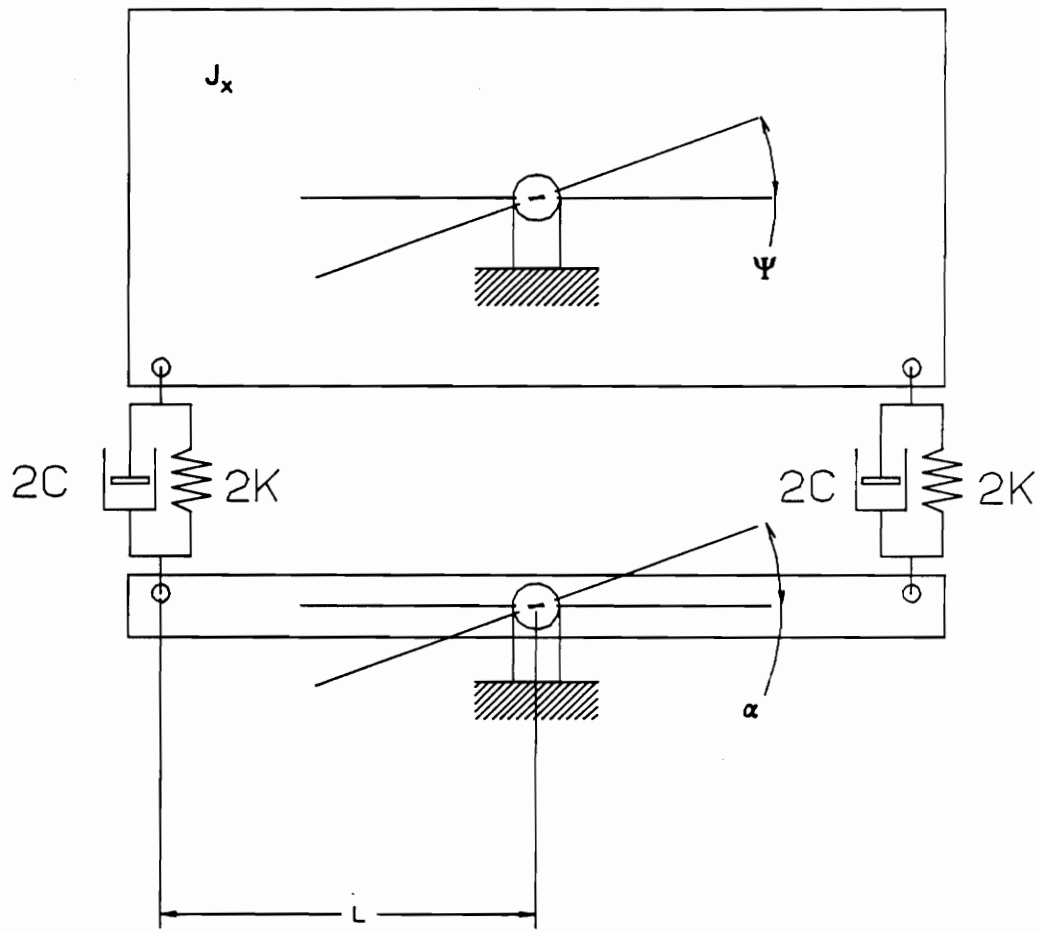


Figure 38. Single degree-of-freedom model of the pitch or roll mode

6.3.3 Parameters to Identify for the System

Under the assumptions in Section 6.3.1, the pitch and roll modes do not provide any more suspension information than the bounce mode. The only additional information from the pitch and roll modes are their moments of inertia. Therefore, only the following parameters are to be identified from the three modes:

1. spring stiffness of each suspension element, k.
2. damping constant of each suspension element, c.
3. carbody mass, m.
4. carbody pitch moment of inertia, J_x.
5. carbody roll moment of inertia, J_z.

The parameters l and d_z were determined from TTC drawings of the actuator locations.

6.4 *Identifiability of the Rail Vehicle Model*

The bounce, pitch, and roll models all contain the same difficulty with identifying parameters as the displacement-input model in Chapter 4. In each mode, only two of the three parameters may be identified when the input is taken as a displacement. One parameter must be determined from separate data prior to the others; identifying three parameters from a single test would give incorrect results. For the single degree-of-freedom system, the identifiability problem is apparent. However, for a more complex system (e.g., equation 6.3.1), the identifiability problem is easily overlooked.

Thus, some of the parameters must be determined using an alternative method. The inertia terms are the easiest parameters to obtain in alternative ways. The next section provides estimates of the three inertia terms.

6.5 *A priori Parameter Information*

The inertia values (m, J_x, and J_z) must be determined prior to the estimation process. Information on these parameters was obtained from

1. Transportation Test Center data.
2. Parameter estimation of inertia terms alone.

6.5.1 Transportation Test Center Data

A paper by Wormley and Tombers (1983) and personal correspondence with engineers at the TTC (Wilson, 1987a,b) provided estimates of the system parameters. Table 10 lists the parameter values obtained from research performed at the TTC.

6.5.2 Parameter Estimation of Inertia Terms

The inertia terms for the entire vehicle were also obtained by reformulating the system models from displacement-input to force-input. Here, the suspension

Table 10. Inertia Values from TTC

Parameter	*Units*	*Value*
carbody mass	$\frac{\text{lb s}^2}{\text{in}}$	41.2
unsprung mass per wheelset	$\frac{\text{lb s}^2}{\text{in}}$	7.25
carbody pitch moment of inertia	lb s² in	2.5×10^6
carbody roll moment of inertia	lb s² in	2.5×10^4
wheelset roll moment of inertia	lb s² in	2.6×10^3
wheelset spin moment of inertia	lb s² in	1.5×10^3

elements simply transmit the force and therefore do not appear in the equations of motion. The equation of motion for the bounce mode is

$$M_v \ddot{Y} = F \qquad [6.5.1]$$

where

M_v = entire vehicle mass (carbody + wheelsets + suspension)

F = force input from all of the actuators

A least-squares parameter estimation was performed for this system. Figure 39 shows the experimental data and the theoretical curve after the performance index converged. Figure 40 and Figure 41 show the convergence of the mass and the performance index, respectively. The mass of the entire vehicle was estimated to be 56.1 $\frac{lb\ s^2}{in}$. The mass of the carbody was then calculated by

$$m = M_v - 2\,m_w = 56.1 - 2(7.25) = 41.6\,\frac{lb\ s^2}{in} \qquad [6.5.2]$$

where

m = carbody mass

m_w = unsprung wheelset and suspension mass

The TTC determined mass and the estimated mass differ by only one percent.

Unfortunately, the pitch and roll mode estimations do not provide results compatible with the TTC values. The estimation of the pitch and roll moments of inertia return values more than 2 times greater than the TTC inertias. This divergence results from the difference between the moment of inertia for the entire vehicle and the moment of inertia for the carbody alone. Subtracting the wheelset inertias from the total inertias compensates some, but still leaves a significant difference. Because of this difficulty, the TTC values were used as the correct values for the carbody inertias.

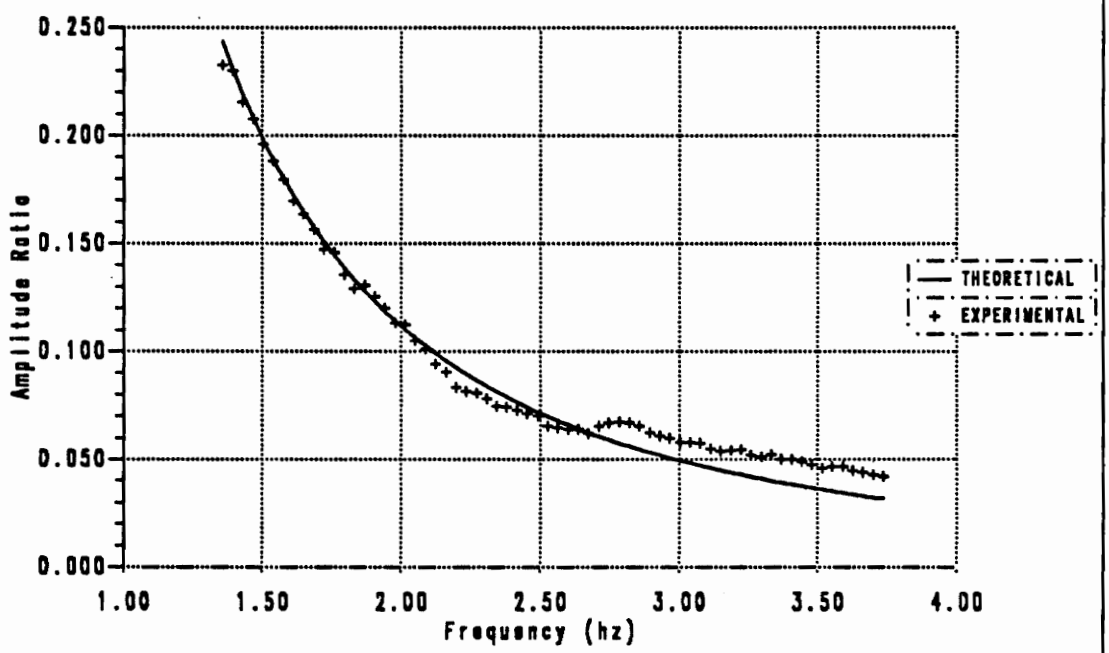


Figure 39. Estimation of total vehicle mass

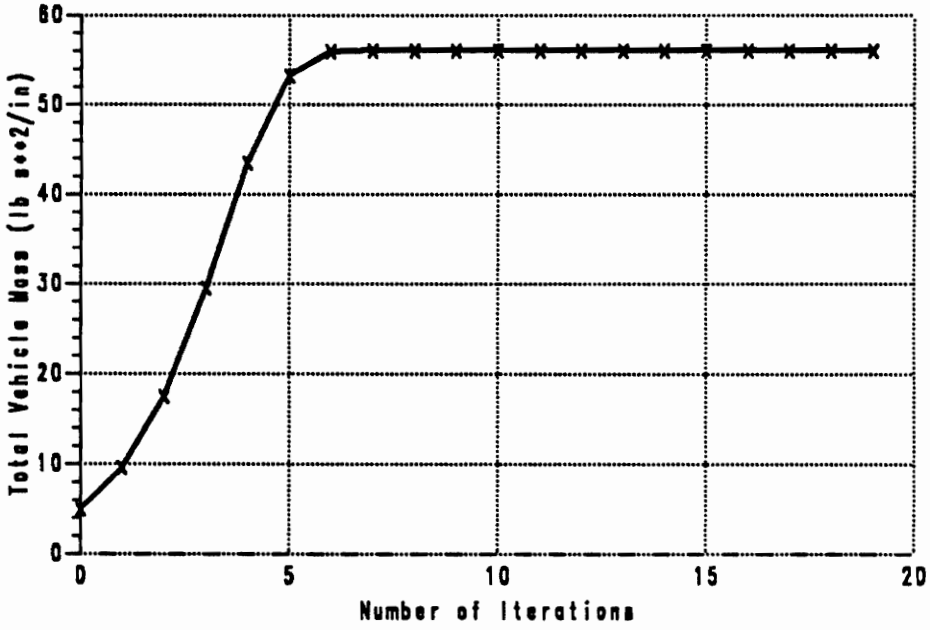


Figure 40. Convergence of mass parameter

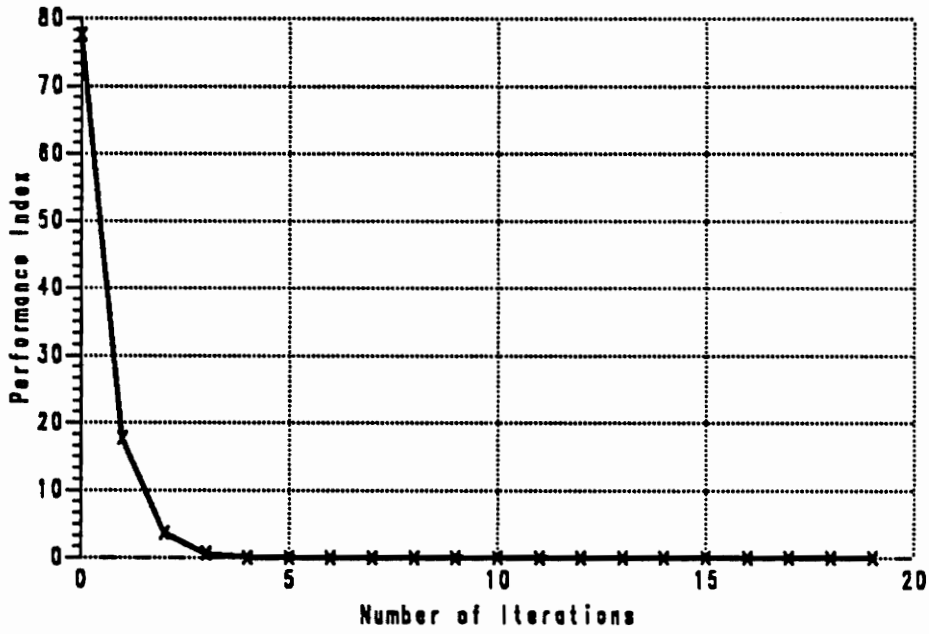


Figure 41. Convergence of performance index

6.6 *Model Estimation*

With one vehicle parameter determined a priori, identifiability problems will not affect the the parameter estimation results. The specific estimation method can now determine the optimum parameters for the system under the assumptions made in Section 6.3.1. Because the roll and pitch moments of inertia were determined from TTC data and the suspension was assumed symmetric, only bounce mode estimation was performed.

6.6.1 Program Structure

Estimation of the vehicle suspension parameters was accomplished by adjusting the model parameters until the experimental and theoretical FRFs matched in a weighted least-squares sense. The performance index minimized was

$$J = \left(\hat{\underline{H}} - \underline{H}\right)' \mathbf{W} \left(\hat{\underline{H}} - \underline{H}\right) + (\underline{\theta}_o - \underline{\theta})' \mathbf{S}_{rr}^{-1} (\underline{\theta}_o - \underline{\theta}) \qquad [6.6.1]$$

where

$\hat{\underline{H}}$ = experimental FRF vector

$\underline{H}$ = theoretical FRF vector

$\mathbf{W}$ = weighting matrix

$\underline{\theta}_o$ = a priori parameter estimate vector

$\underline{\theta}$ = current parameter estimate vector and

$\mathbf{S}_{rr}$ = parameter error covariance matrix

The first term of equation [6.6.1] is the standard weighted least-squares performance index. The experimental FRF vector, $\hat{\underline{H}}$, is a function of frequency only and the theoretical FRF vector, $\underline{H}$, is a function of both frequency and the current parameter vector, $\underline{\theta}$. Several different weighting matrices, **W**, were used to determine their effect on the convergence of the parameter estimation.

The second term of equation [6.6.1], called the Bayesian term, influences changes in the parameter vector, $\underline{\theta}$, from the initial parameter vector, $\underline{\theta}_o$. The Bayesian term can impact the estimation convergence greatly. For example, when the elements of the parameter error covariance matrix, $\mathbf{S}_{rr}$, approach zero, indicating a perfect initial guess, the Bayesian term dominates the performance index and does not allow any change in the parameter vector. However, when the elements of $\mathbf{S}_{rr}$ are very large, the Bayesian term has no influence and the estimation degenerates to a weighted least-squares class estimation.

Determining a first guess for the individual terms of the parameter error covariance matrix is often difficult. The individual terms of $\mathbf{S}_{rr}$ are equal to the square of the estimated standard deviation of the parameter. Thus, the parameter error covariance is

$$\mathbf{S}_{rr_{ii}} = [\eta\, \theta_{o_i}]^2 \qquad [6.6.2]$$

where

i = 1, 2, ..., number of parameters

η = percentage of parameter for confidence of plus or minus one standard deviation ($\pm \sigma$)

θ_{o_i} = initial guess of the i^{th} parameter

Efficient use of the parameter error covariance matrix requires good a priori information on the confidence of the initial estimates.

The parameter estimation of the vehicle suspension used the following equations repeated here from Chapter 2.

$$\underline{\theta}^{\dagger} = \underline{\theta} + \left[\mathbf{T}'\,\mathbf{W}\,\mathbf{T} + \mathbf{S}_{rr}^{-1}\right]^{-1}\left[\mathbf{T}'\,\mathbf{W}\,(\hat{\underline{H}} - \underline{H}) + \mathbf{S}_{rr}^{-1}(\underline{\theta}_o - \underline{\theta})\right] \qquad [6.6.3]$$

$$\mathbf{S}_{rr}^{\dagger} = \left(\mathbf{T}'\,\mathbf{W}\,\mathbf{T} + \mathbf{S}_{rr}^{-1}\right)^{-1} \qquad [6.6.4]$$

where

$\underline{\theta}^{\dagger}$ = revised estimate of $\underline{\theta}$

$\mathbf{S}_{rr}^{\dagger}$ = revised estimate of $\mathbf{S}_{rr}$

and

$$\mathbf{T} = \frac{\partial \underline{H}}{\partial \underline{\theta}}$$

The estimation process then followed the steps below (Fries, 1983):

1. Set $\underline{\theta} = \underline{\theta}_o$.
2. Compute $\underline{\theta}^{\dagger}$.
3. Let $\underline{\theta} = \underline{\theta}^{\dagger}$ and compute a new $\underline{\theta}^{\dagger}$.
4. Repeat step 3 until $\underline{\theta}^{\dagger}$ converges.
5. Compute $\mathbf{S}_{rr}^{\dagger}$.

This estimation always converged when using the $\mathbf{S}_{rr}$ term but did not always converge without it. The final values of $\mathbf{S}_{rr}^{\dagger}$ provide one indication of the quality of the final parameters returned from the estimation. A computer software package, Matlab (Matlab User's Guide, 1987) was used to implement this algorithm on an IBM-AT style personal computer.

6.6.2 Estimation of Bounce Model

6.6.2.1 Frequency Response Function Used

The FRF used for the estimation of the suspension properties, $\hat{H}$, was computed from averaged input and output measurements. The input signal was defined as

$$z = \frac{1}{4}(1AZ + 1BZ + 3AZ + 3BZ) \qquad [6.6.5]$$

where

1AZ = input displacement from transducer 1AZ

1BZ = input displacement from transducer 1BZ

3AZ = input displacement from transducer 3AZ

3BZ = input displacement from transducer 3BZ

The output signal was defined as

$$Y = \frac{1}{4}(D15Z + D16Z + D18Z + D19Z) \qquad [6.6.6]$$

where

D15Z = output displacement from transducer D15Z

D16Z = output displacement from transducer D16Z

D18Z = output displacement from transducer D18Z

D19Z = output displacement from transducer D19Z

The experimental FRF of the averaged input and output signals is shown in Figure 42. Figure 25 shows the location of each of the transducers.

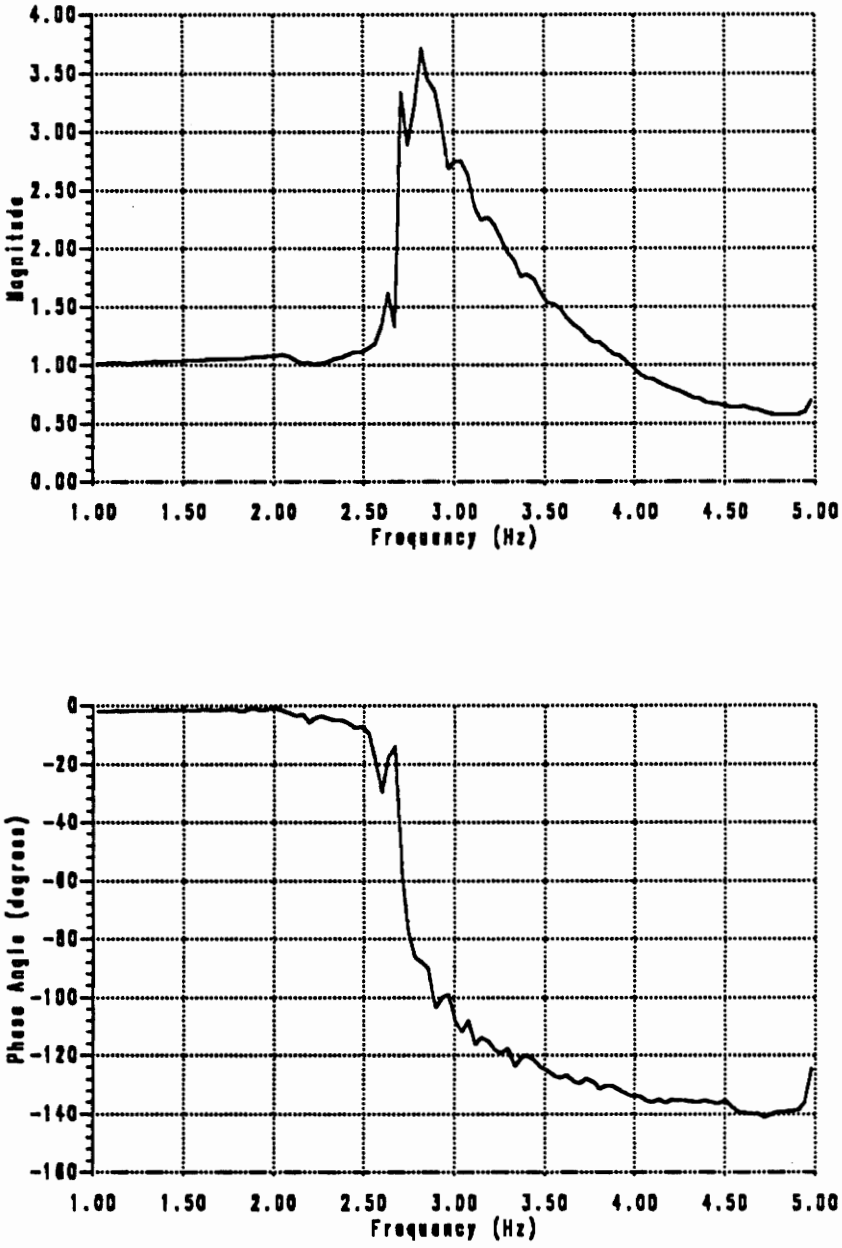


Figure 42. Frequency response function of averaged bounce response

6.6.2.2 *Weighting Functions*

The estimation was first performed with each term of $\mathbf{S}_{rr}$ set to 10^{50}. This allowed the estimation to proceed with virtually no constraint on the parameter values. As a result, the estimation did not always converge, or converged to physically unreasonable values. These difficulties do not result from identifiability problems but from an inadequate weighting function or from incorrect modeling of the real dynamics. Next, the Bayesian term was added to some of the estimations.

Several different weighting functions were investigated to determine their effect on the estimation. The weighting functions investigated were:

1. Observation error covariance matrix, $\mathbf{S}_{\varepsilon\varepsilon}^{-1}$.
2. Identity matrix, $\mathbf{I}$.
3. Matrix with experimental FRF as its diagonal elements, $\mathbf{W}_H$.
4. Matrix with numerous zeros on the diagonal, removing specific frequencies from the estimation, $\mathbf{W}_0$.

6.6.2.3 *Estimation Using Observation Error Covariance Matrix*

The inverse observation error covariance matrix, $\mathbf{S}_{\varepsilon\varepsilon}^{-1}$, is theoretically the best weighting function to use in an estimation process. Chapter 2 contains the theoretical development of this weighting function. The diagonal terms of $\mathbf{S}_{\varepsilon\varepsilon}$ are estimates of the FRF error at each frequency. Bendat and Piersol (1986) included an equation for estimating the variance of the FRF error. For the single-input single-output case, the equation is

$$r^2(f) = \frac{1}{n_{rec} - 1} F_{n_1 n_2 \alpha} \left[1 - \gamma^2_{yx}(f)\right] \frac{G_{yy}(f)}{G_{xx}(f)} \qquad [6.6.7]$$

where

$r^2(f) \geq |\hat{H} - H|^2$

n_{rec} = number of ensemble averaged segments

$F_{n_1 n_2 \alpha}$ = 100α percentage point of an F distribution with $n_1 = 2$ and $n_2 = (2\, n_{rec} - 2)$ degrees of freedom

$\gamma^2_{yx}(f)$ = coherence between the output and the input

$G_{yy}(f)$ = PSD of the output

$G_{xx}(f)$ = PSD of the input

The coherence function is computed by

$$\gamma^2_{xy}(f) = \frac{|G_{xy}(f)|^2}{G_{xx}(f)\, G_{yy}(f)} \qquad [6.6.8]$$

The elements of the observation error covariance matrix are given by

$$\mathbf{S}_{\varepsilon\varepsilon_{jj}} = r_j^2 \qquad [6.6.9]$$

The parameters of the F distribution provide an error band of plus or minus one standard deviation ($\pm 1\sigma$). The measurement errors are assumed to be uncorrelated; therefore $\mathbf{S}_{\varepsilon\varepsilon}$ is diagonal. The weighting matrix, **W**, used in equations 6.6.3 and 6.6.4 is the inverse of the observation error covariance matrix, or

$$\mathbf{W}_{ii} = \frac{1}{\mathbf{S}_{\varepsilon\varepsilon_{ii}}} \qquad [6.6.10]$$

Figure 43 shows the coherence, γ^2_{xy} , of the bounce data FRF. Bendat and Piersol provided distribution tables for values of $F_{n_1 n_2 \alpha}$. Using this data, the diagonal values of $S_{\alpha\alpha}$ are computed and shown in Figure 44. Figure 45 shows the weighting function.

Though the observation error covariance matrix is theoretically sound, several practical difficulties hamper its effectiveness. From equation [6.6.7], the covariance of the data approaches zero as the coherence approaches one. This indicates when no measurement error exists, the reliability of the data is perfect and therefore those frequencies are weighted heavily in the estimation. Practically, as Figure 45 shows, some sections of the data can be weighted much more heavily than the rest of the data.

The estimation using the observation error covariance matrix was performed first without the Bayesian terms ($S_{rr_{ii}} = 10^{50}$) . Next, the estimation included a Bayesian term computed from guesses of the initial parameter standard deviation. The magnitude of the Bayesian elements were equal to

$$S_{rr_{ii}} = [\eta \, \theta_{o_i}]^2 \qquad [6.6.11]$$

where

$$\eta = 0.20$$

For both the estimations, the heavy weighting of the FRF at low frequencies and the deemphasis of the FRF near the resonant peak caused a poor estimation of the parameters for this system. Figure 46 shows the final estimate of the FRF for estimates without a Bayesian term (PEST8) and for estimates containing a Bayesian term (PEST12). Table 11 and Table 12 show the final estimates of the stiffness and damping compared to values obtained using other weighting functions.

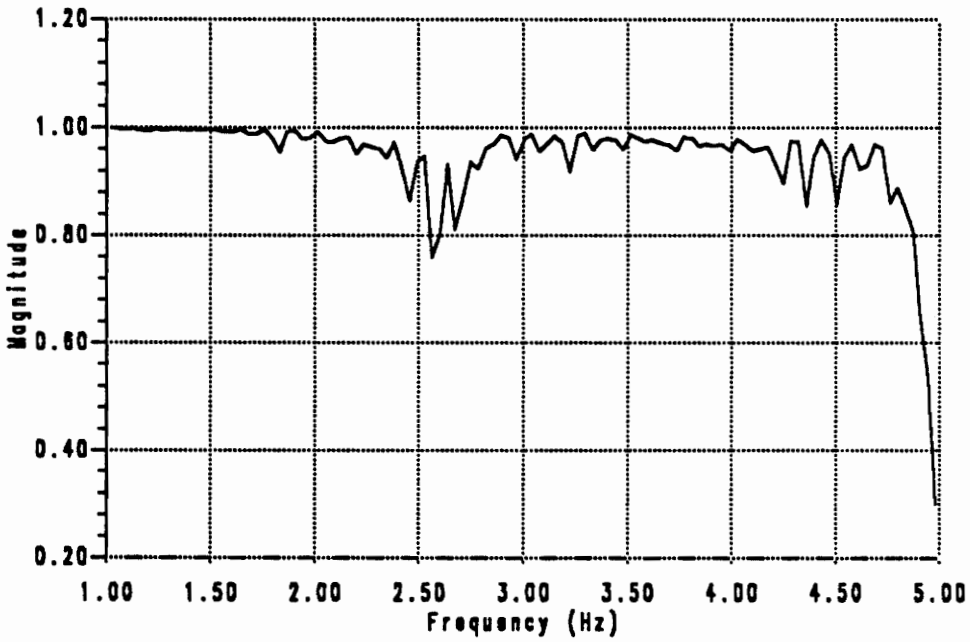


Figure 43. Coherence function for the averaged bounce FRF

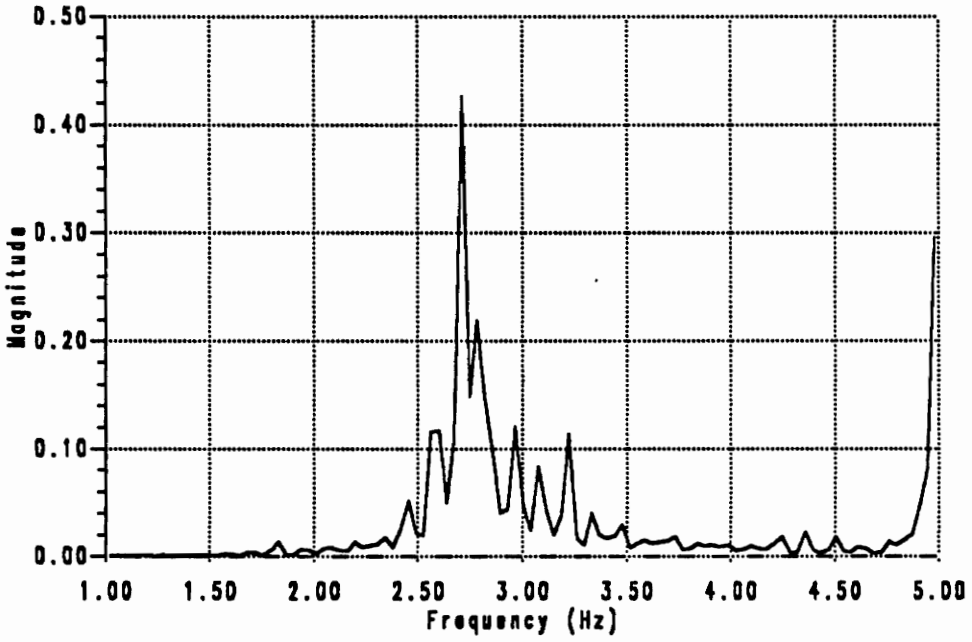


Figure 44. Observation error covariance for the averaged bounce FRF

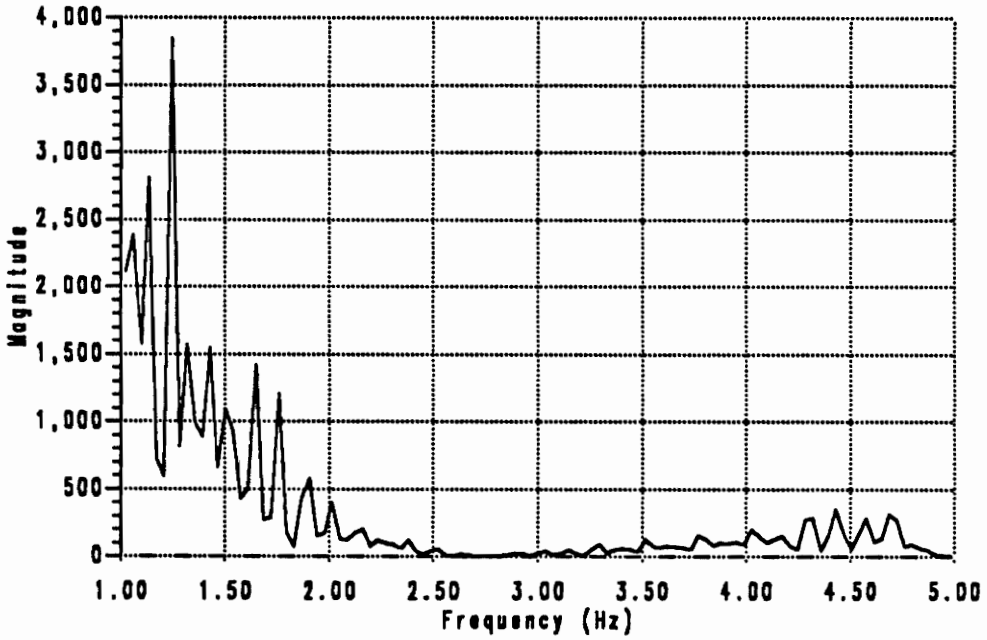


Figure 45. Weighting function for the averaged bounce FRF (S_{tt}^{-1})

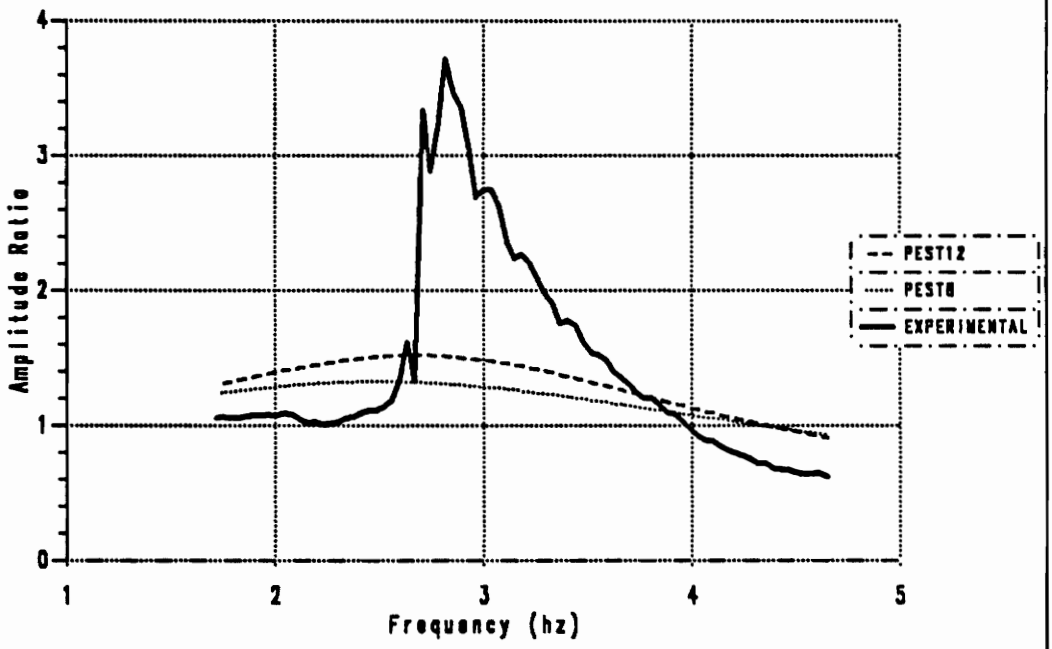


Figure 46. Final estimates of the FRF using observation error covariance weighting

Table 11. Parameter estimation results for stiffness

Method Label	*Weighting Function*	*Stiffness Initial Guess (lb/in)*	*Initial Parameter Error Covariance*	*Stiffness Final Value (lb/in)*	*Final $\pm 1\sigma$ Error (percent)*
PEST1	$[W_0\ W_{H1}]$	12000	10^{50}	14100	8.0
PEST3	$[W_0\ I]$	12000	10^{50}	14700	13
PEST5	$[W_0\ S_{\varepsilon\varepsilon}^{-1}]$	12000	10^{50}	50600	165
PEST7	W_{H4}	12000	10^{50}	13000	1.1
PEST8	$S_{\varepsilon\varepsilon}^{-1}$	12000	10^{50}	15300	3.0
PEST9	I	12000	5.76 E06	14300	5.0
PEST10	I	12000	10^{50}	15200	7.6
PEST11	W_{H1}	12000	10^{50}	14400	4.7
PEST12	$S_{\varepsilon\varepsilon}^{-1}$	12000	5.76 E06	15400	1.8

Table 12. Parameter estimation results for damping

Method Label	*Weighting Function*	*Damping Initial Guess (lb s / in)*	*Initial Parameter Error Covariance*	*Damping Final Value (lb s / in)*	*Final ± 1σ Error (percent)*
PEST1	$[W_0\, W_{H^1}]$	250	10^{50}	257	8.7
PEST3	$[W_0\, I]$	250	10^{50}	284	15
PEST5	$[W_0 S_{zz}^{-1}]$	250	10^{50}	4346	79
PEST7	W_{H^4}	200	10^{50}	227	1.6
PEST8	S_{zz}^{-1}	250	10^{50}	979	4.0
PEST9	I	200	1600	287	8.6
PEST10	I	200	10^{50}	373	13
PEST11	W_{H^1}	200	10^{50}	298	7.6
PEST12	S_{zz}^{-1}	250	2500	727	2.9

Without the Bayesian term, both parameter values increased to unrealistic values. The weighting of the observation error covariance matrix dominated the performance index at low frequencies. Therefore, the estimation converged primarily to minimize the error at low frequencies. Even with a strong Bayesian term, the observation error term dominated the performance index. The stiffness parameter is helped significantly by the addition of the Bayesian term. However, the observation error weighting still dominated the damping term.

6.6.2.4 Estimation Using Identity Matrix

The next weighting function incorporated the identity matrix into the formulation. For this case,

$$\mathbf{W} = \mathbf{I} \qquad [6.6.12]$$

and the estimation decomposed to a standard unweighted least-squares class estimation. This estimation method was used in the test problem of Chapter 4. The estimation using the identity matrix was performed both with and without the Bayesian term. The magnitude of the Bayesian terms were equal to

$$\mathbf{S}_{rr_{ii}} = [\eta\, \theta_{o_i}]^2 \qquad [6.6.13]$$

where

$$\eta = 0.20$$

Figure 47 shows the final estimated FRF for the estimation with the Bayesian term (PEST9) and without the Bayesian term (PEST10). The response using the Bayesian

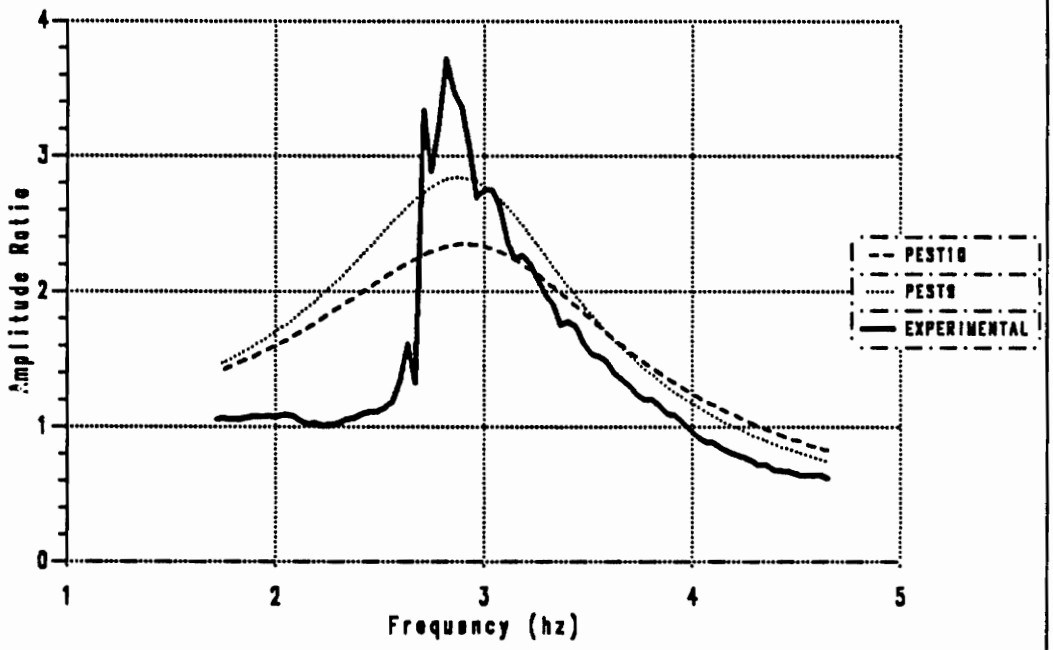


Figure 47. Final estimates of the FRF using identity matrix weighting

term provides a better estimate of the peak value, although neither of the estimates closely model the experimental response.

6.6.2.5 *Estimation Using Experimental FRF Weighting*

An estimation which used powers of the experimental FRF as the diagonal elements of the weighting function was the most successful of all the estimations. The Bayesian term was not included in this estimation. The weighting function of the estimation is given by

$$W_{H_{ii}^n} = \left(\hat{H}_i\right)^n \qquad [6.6.14]$$

where

$\hat{H}_i$ = i^{th} value of the experimental FRF vector

n = integer power

For these estimation tests, two different powers of the FRF were used: H^1 and H^4. Figure 48 shows the final estimated FRF with both of these weightings. The final stiffness and damping values are listed in Table 11 and Table 12. PEST11 uses the W_{H^1} weighting and PEST7 uses the W_{H^4} weighting.

PEST7 provides the closest estimate of the peak value of the FRF of any of the estimation methods. This method provides a good estimate because the values of the FRF near resonance are weighted much more heavily than estimates far from the resonant peak.

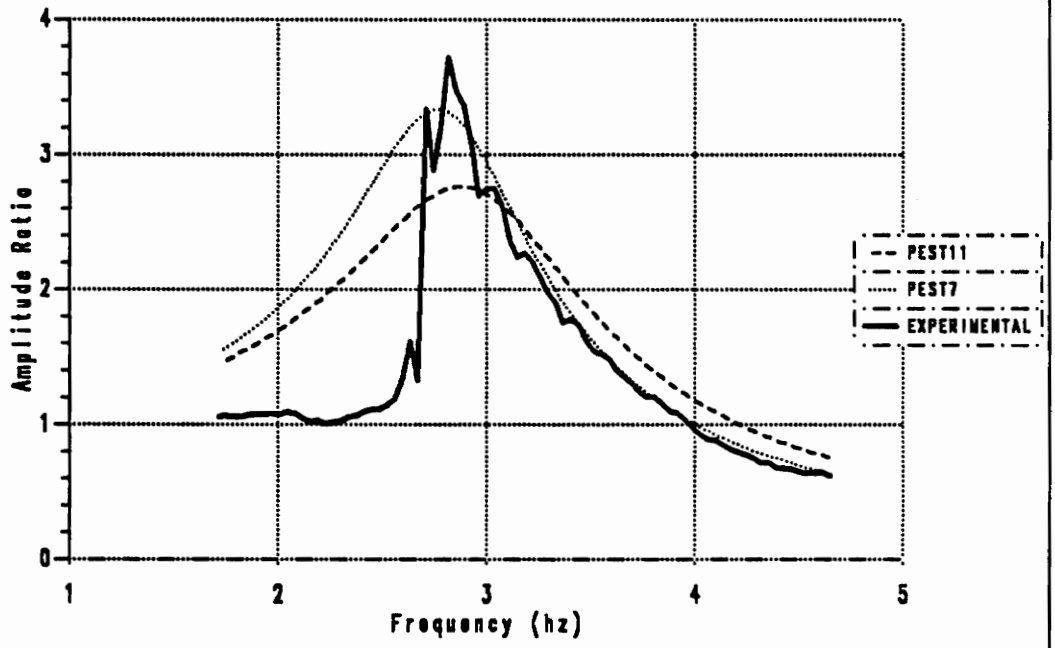


Figure 48. Final estimates of the FRF using experimental FRF weighting matrix

6.6.2.6 *Estimation Using Selected Frequencies*

A final weighting function incorporated only selected frequencies from the experimental FRF. This method, called "selection weighting" in this thesis, combines two weighting matrices. The first matrix is identical to one of the weighting functions above. The second matrix contains zeros on the diagonal location of each frequency excluded from the estimation. Therefore, the weighting matrix becomes

$$\mathbf{W} = \mathbf{W}_n \mathbf{W}_0 \qquad [6.6.15]$$

where

$\mathbf{W}_n$ = normal weighting function ($\mathbf{S}_{\alpha\alpha}^{-1}$, $\mathbf{I}$, $\mathbf{W}_{H^n}$)

$\mathbf{W}_0$ = matrix with zeros at locations to exclude frequencies

Seventeen frequencies were used in the estimation and are shown in Figure 49. The frequencies selected are highlighted. The matrix $\mathbf{W}_0$ must also be included in the second term of equations 6.6.1 and 6.6.3 to maintain FRF vectors of equal length. The selection weighting method was used in conjunction with the following other weighting functions:

1. Inverse error covariance weighting, $\mathbf{S}_{\alpha\alpha}^{-1}$.
2. Identity matrix, $\mathbf{I}$.
3. Matrix with FRF used in estimation as its diagonal elements, $\mathbf{W}_H$.

The Bayesian term was not used in the estimation of any of these tests. The FRFs obtained from this estimation are shown in Figure 50. The final stiffness and damping values are listed in Table 11 and Table 12. Once again, the estimation using $\mathbf{S}_{\alpha\alpha}$ showed the worst results. Here, the stiffness and damping parameters increased to values so large that the FRF appeared as a horizontal line in the test

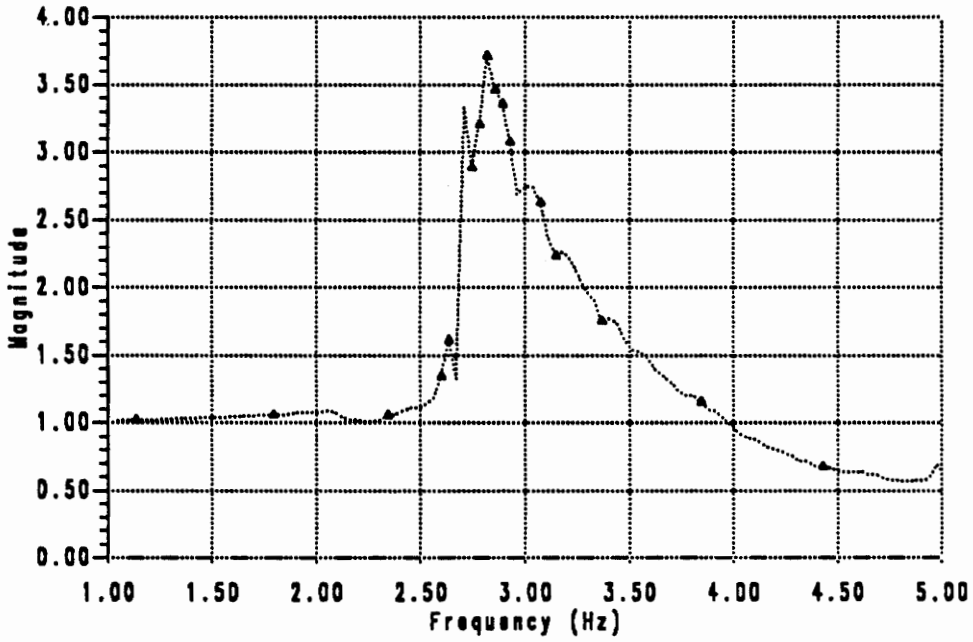


Figure 49. Frequencies selected for the estimation

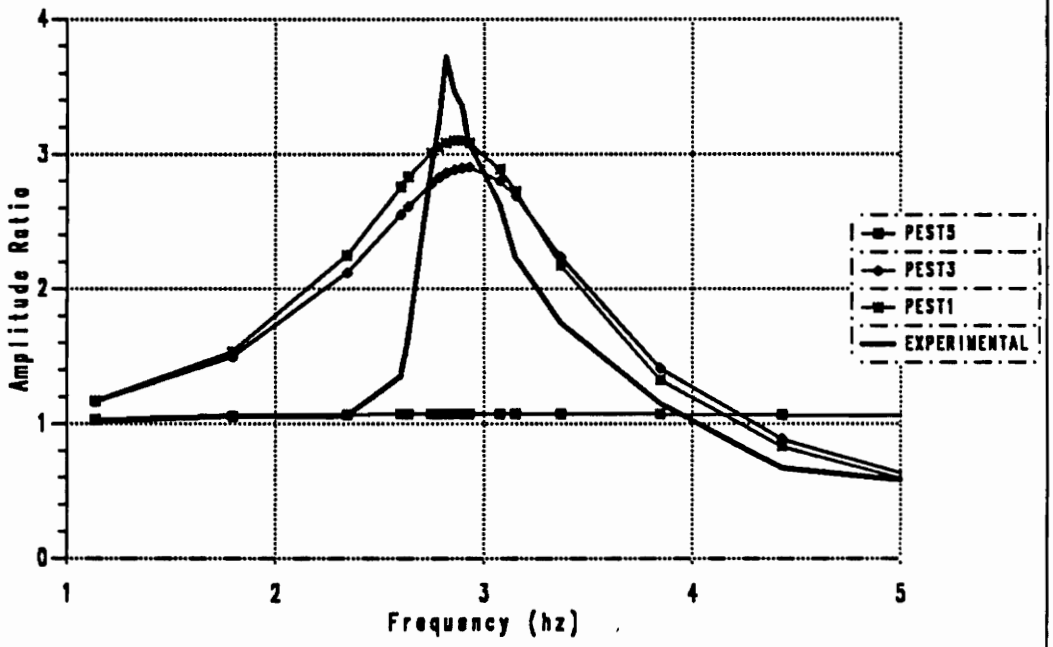


Figure 50. Comparison of different weighting functions using selection weighting method

frequency range. The best results from the selection weighting method were obtained using the weighting matrix W_{H1}.

6.7 Conclusions

The estimation results from the previous section varied widely depending on the weighting function. Figure 51 and Figure 52 show the results of all of the estimations except PEST5 which is clearly in error (see Figure 50). Figure 50 and Figure 53 show the results from all of the estimations. Unfortunately, no estimation provided good agreement throughout the entire frequency range. Only one estimation, using W_{H4} as the weighting function, predicted the experimental peak value within 10 percent.

The reliability of the estimation is limited by the accuracy of the dynamic model to predict the real behavior of the system. The linear and symmetric assumptions made in section 6.3.1 limit the validity of the results. Several nonlinearities are suspected in the system. The dominant nonlinearity is probably Coulomb friction in the leaf spring suspension. The Coulomb friction nonlinearities are seen from displacement transducers and accelerometers mounted on the vehicle. Figure 54 and Figure 55 show representative displacements and accelerations of the rail vehicle during the bounce test. The displacement record shows a slight aberration from a pure sine wave near the peaks of the signal. The acceleration record accentuates this variation. These two records indicate a stick-slip condition exists in the vehicle suspension.

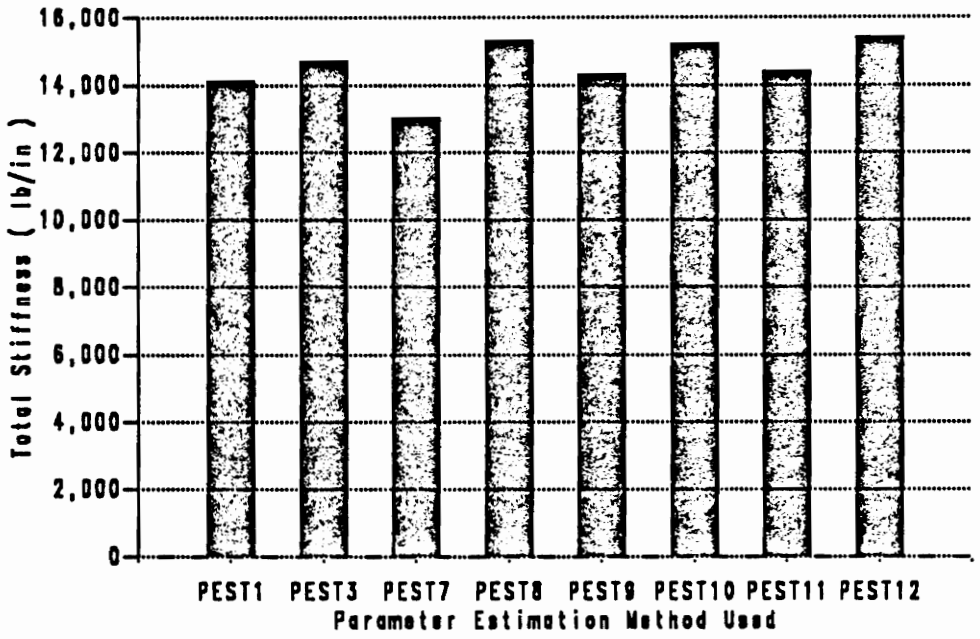


Figure 51. Final estimates of the stiffness parameter from all estimations (except PEST5)

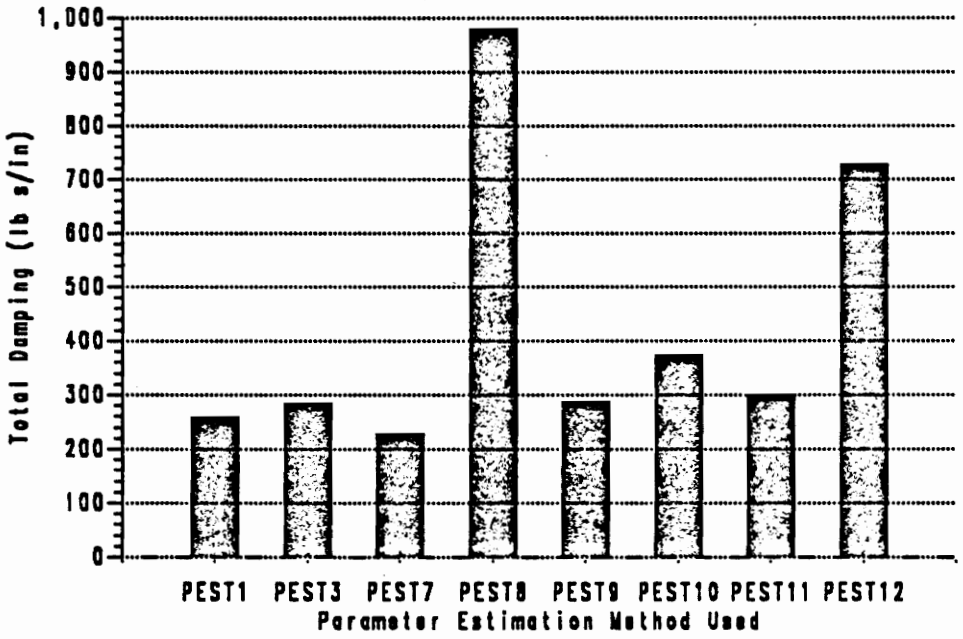


Figure 52. Final estimates of the damping parameter from all estimations (except PEST5)

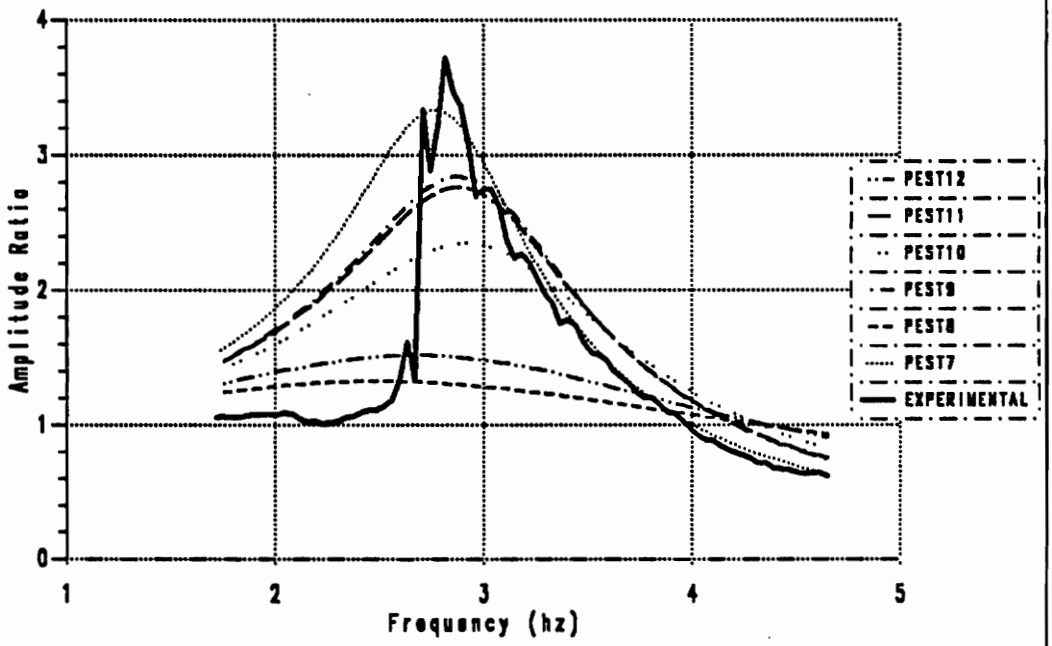


Figure 53. Final estimates of FRF magnitudes using all weighting functions

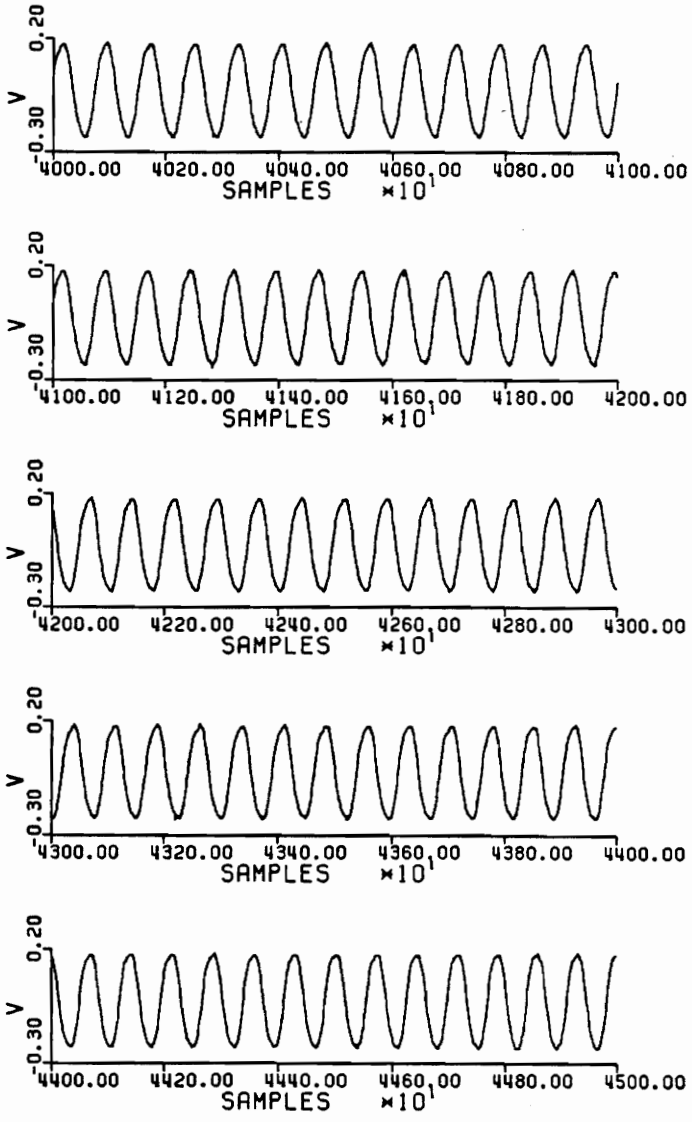


Figure 54. Displacement time response from rail vehicle

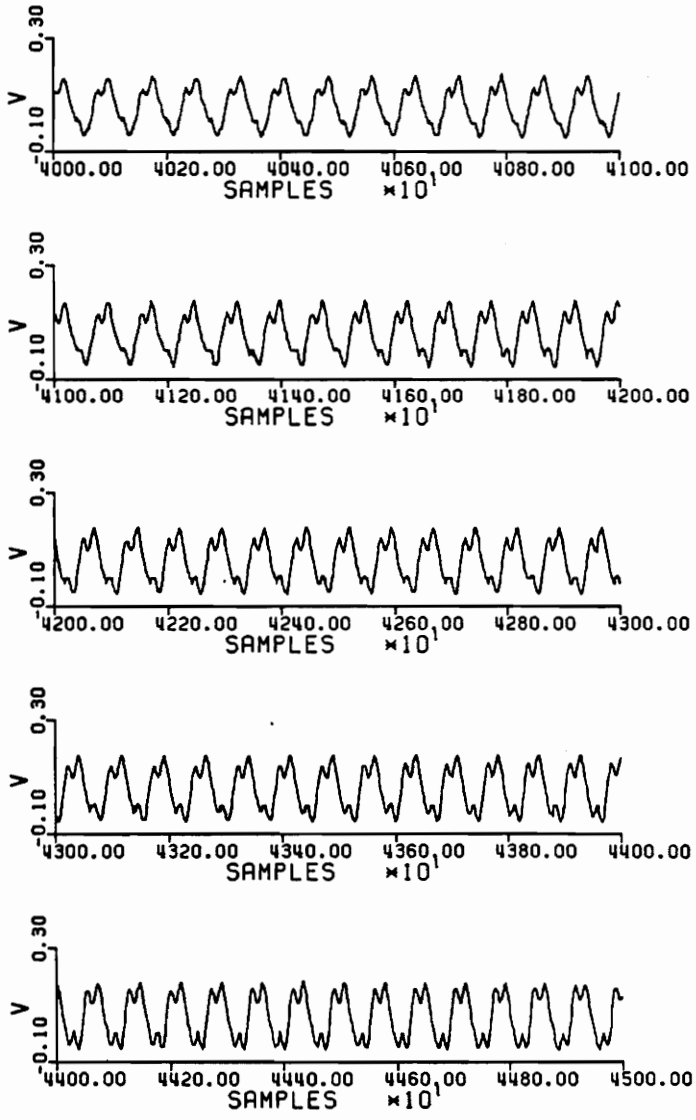


Figure 55. Acceleration time response from rail vehicle

The stick-slip condition probably contributes to the steepness of the FRF between 2.5 and 2.8 Hz. This region was the most difficult to match with a linear estimation. The steepness of the response indicates the vehicle translates with the input at low frequencies until the friction in the leaf spring is overcome and the carbody breaks free from the suspension. At higher frequencies, a linear second-order model fits the response closely. For example, after 3 Hz the estimation using W_{H4} provides a good estimate of the FRF.

The peak FRF amplitude and its frequency are important measures of ride quality and stability for rail vehicles. Thus, the method which most accurately predicts these characteristics is used to provide the final parameter values for the vehicle. Thus, the estimation method using W_{H4} was chosen to predict the parameters for the vehicle. The final estimated parameters for the vehicle are presented in Table 13.

Table 13. Final parameter values estimated for the rail vehicle suspension

Parameter	*Units*	*Value*
suspension stiffness per group	$\frac{\text{lb}}{\text{in}}$	3250
suspension damping per group	$\frac{\text{lb s}}{\text{in}}$	56.8

Chapter 7

Summary, Conclusions, and Recommendations

7.1 Summary

This thesis provides an investigation into basic linear estimation techniques applicable to rail vehicle dynamics. A modified least-squares algorithm was developed to determine the best-fit model of a dynamic system to experimental frequency response functions. The addition of a Bayesian term in the performance index accounted for a priori knowledge of the parameters.

Two methods were presented to determine the identifiability of linear state-space dynamic models. These methods were later used to check the identifiability of two systems, only one being identifiable. Next, the interaction between identifiability and parameter estimation was investigated. A test case showed how identifiability problems can be masked by the parameter estimation.

The vehicle investigated was a lightweight intermodal car with single axle suspension. Engineers at the Transportation Test Center performed all of the experiments and provided data for this thesis.

Standard data processing techniques were used to compute frequency response functions for the vehicle. From the frequency response functions and a hypothesized linear model, the parameters of the rail vehicle were estimated. Nine different combinations of weighting functions and Bayesian terms were used in the performance index. These combinations significantly affected the final parameter values.

7.2 Conclusions

The output error parameter estimation technique is well suited for linear rail vehicle dynamics. The method requires a minimal amount of data to perform the estimation and is effective in the presence of measurement error. A disadvantage of the method is its sensitivity to process or modeling errors.

The identifiability of a system impacts parameter estimation results significantly. When the system is nonidentifiable, parameter estimation results can return values which are clearly in error, or can return values which seem reasonable on first inspection. In both cases, however, the parameters do not uniquely define the system. Also, the Bayesian term must be used carefully. A strong Bayesian term in the presence of an identifiability problem further occludes the results of the estimation. Often, false confidence can be instilled in estimation results if the Bayesian term dominates the performance index.

Identifiability problems are often difficult to recognize in dynamic systems. Even in low order, single degree-of-freedom systems, the identifiability is easily overlooked. Two methods, one developed by Reid and the other developed by Grewal and Glover, provide algorithms for determining the identifiability of linear state-space models. These methods, if implemented before experimentation, can reduce the likelihood of estimation difficulties met in nonidentifiable systems.

Parameter estimation results vary widely depending on the weighting function used. The theoretically sound measurement error covariance weighting matrix contains several practical difficulties and here provided worse results than the other weighting functions. The best weighting functions for this work incorporated the experimental FRF values into their formulation. This weighting function resulted in a damped natural frequency within one percent and a peak value within ten percent of the experimental response.

Finally, the reliability of the estimation is limited by the accuracy of the dynamic model. A linear spring-mass-damper system oversimplifies the dynamics of this rail vehicle. As a result, no estimation provided good magnitude convergence throughout the frequency range. While most of the estimations accurately predicted the system resonance frequency, only one estimation predicted the experimental peak value within 10 percent. For this rail vehicle, Coulomb friction is probably the dominant nonlinearity.

7.3 Recommendations for Future Work

The estimation results of this thesis were limited by the linear and symmetric assumptions. The nonlinearity suspected to influence the response the most is Coulomb (dry) friction in the vehicle suspension. Therefore, a logical continuation of this work would estimate a more complex dynamic model.

To incorporate this suggestion, nonlinear parameter estimation techniques applicable to rail vehicle dynamics must be developed. A variation of the describing function approach to nonlinear systems might be used to formulate a new estimation algorithm.

Unfortunately, identifiability methods for nonlinear systems are rare. The identifiability methods presented in Chapter 3 cannot be applied to systems with friction nonlinearities. Much research needs to be completed on the nonlinear system identifiability to give engineers and mathematicians tools such as Reid's method and Grewal and Glover's method provide for linear systems.

References

Beck, J. V., and K. J. Arnold, *Parameter Estimation in Science and Engineering*, John Wiley and Sons, New York, 1977.

Bendat, J. S., and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, John Wiley and Sons, New York, 1986.

Broersen, P. M. T., "Estimation of multivariable railway vehicle dynamics from normal operating records," *Identification and System Parameter Estimation*, P. Eykhoff ed., American Elsivier Publishing Company, New York, 1973, pp. 425-434.

Chen, C. T., *Introduction to Linear System Theory*, Holt, Rinehart and Winston, Inc., New York, 1970.

Cobelli, C., A. Lepschy, and G. Romanin-Jacur, "Identifiability of compartmental systems and related structural properties," *Mathematical Biosciences*, Vol. 44, 1979, pp. 1-18.

Crossley, T. R., and B. Porter, "Eigenvalue and eigenvector sensitivities in linear systems theory," *International Journal of Control*, Vol. 10, No. 2, 1969, pp. 163-170.

Delforge, J., "Letter to the Editor," *Mathematical Biosciences*, Vol. 61, 1982, pp. 299-305.

Delforge, J., "Local identifiability of linear systems in a general input-output configuration," *International Journal of Systems Sciences*, Vol. 16, No. 8, 1985, pp. 1015-1026.

Delforge, J., "Necessary and sufficient structural condition for local of a system with linear compartments," *Mathematical Biosciences*, Vol. 54, 1981, pp. 159-180.

Delforge, J., "New results on the problem of identifiability of a linear system," *Mathematical Biosciences*, Vol. 52, 1980, pp. 73-96.

Delforge, J., "The problem of structural identifiability of a linear compartmental system: solved or not?," *Mathematical Biosciences*, Vol. 36, 1977, pp. 119-125.

Distefano, J. J., and C. Cobelli, "On parameter and structural identifiability: nonunique observability/reconstructability for identifiable systems, other ambiguities and new definitions," *IEEE Trans. Autom. Control*, Vol. 25, No. 4, 1980, pp. 830-833.

Eykhoff, P., *System Identification: Parameter and State Estimation*, John Wiley and Sons, London, 1974.

Fallon, W. L., "An Investigation of Railcar Model Validation," Master's thesis, Arizona State University, Tempe, Arizona, May, 1977.

Forsythe, G. E., M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1977.

Fries, R. H., *Estimation of transit rail vehicle parameters from roller rig tests*, PhD dissertation, Arizona State University, Tempe, Arizona, August, 1983.

Fries, R. H., and B. M. Coffey, "A state-space approach to the synthesis of random vertical and crosslevel rail irregularities," *ASME Paper No. 87-WA/DSC-38*, New York, 1987.

Garg, V. K., and R. V. Dukkipati, *Dynamics of Railway Vehicle Systems*, Academic Press, Ontario, 1984.

Gelfand, I. M., and S. V. Fomin, *Calculus of Variations*, Prentice Hall, Englewood Cliffs, NJ, 1963, p. 11.

Gilan, A., *Parameter identification techniques for rail vehicle dynamic models*, PhD dissertation, Massachusetts Institute of Technology, Cambridge, MA, May, 1981.

Gilan, A., and J. K. Hedrick, "Validating rail vehicle dynamic models: a case study," *ASME Paper No. 82-RT-6*, New York, 1982.

Gostling, R. J., and N. Cooperrider, "Validation of railway vehicle lateral dynamics models," *Vehicle System Dynamics*, Vol. 12, 1983, pp. 179-202.

Grewal, M. S., *Modeling and identification of freeway traffic systems*, PhD dissertation, University of Southern California, Los Angeles, CA, 1974.

Grewal, M. S., and K. Glover, "Identifiability of linear and nonlinear dynamical systems," *IEEE Trans. Autom. Control*, Vol. 21, 1976, pp. 833-837.

Grewal, M. S., and H. J. Payne, "Identification of parameters in a freeway traffic model," *IEEE Trans. Systems, Man, Cybernetics*, Vol. 6, No. 3, 1976, pp. 176-185.

Harris, F. J., "On the use of windows for harmonic analysis with the discrete fourier transform," *Proceedings of the IEEE*, Vol. 66, No. 1, 1978, pp. 51-83.

Hasselman, T. K., and L. Johnson, "Validation and verification of rail vehicle models," *ASME Paper No. 79-WA/DSC-8*, New York, 1979.

Hedrick, J. K., D. N. Wormley, A. K. Kar, W. Murray, and W. Baum, "Performance limits of rail passenger vehicles: conventional radial and innovative trucks," in *DOT-OS-70052*, U. S. Department of Transportation, September 17, 1979.

Hedrick, J. K., D. N. Wormley, A. K. Kar, W. Murray, and W. Baum, "Performance limits of rail passenger vehicles: evaluation and optimization," in *DOT-OS-70052*, U. S. Department of Transportation August 19, 1978.

Hsia, T. C., *System Identification: Least-Squares Methods*, Lexington Books, Lexington, MA, 1977.

Inskeep, D. W., and J. A. Roberts, "Railroad vehicle dynamic testing capability and success," *ASME Paper No. 82-WA/RT-9*, New York, 1982.

Irani, F. D., N. G. Wilson, C. L. Urban, and C. E. Eickhoff, "Developments in testing and analysis for the safety aspects of new vehicle designs," *ASME Paper No. 86-WA/RT-11*, New York, 1986.

Isenberg, J., "Progressing from least-squares to Bayesian estimation," *ASME Paper No. 79-WA/DSC-16*, New York, 1979.

Jain, V. K., and G. J. Dobeck, "System identification techniques: a tutorial review," *ASME Paper No. 79-WA/DSC-20*, New York, 1979.

Junkins, J. L., *An Introduction to Optimal Estimation of Dynamical Systems*, Sijtoff and Noordhoff International Publishers, Alphen aan den Rign, the Netherlands, 1978.

Kallenbach, R. G., "Identification methods for vehicle system dynamics," *Vehicle System Dynamics*, Vol. 16, 1987, pp. 107-127.

Kilpatrick, J. J., *The Writer's Art*, Andrews, McMeel and Parker, Inc., Kansas City, 1984.

Law, E. H., and N. K. Cooperrider, "A survey of railway vehicle dynamics research," *Journal of Dynamic Systems, Measurement, and Control*, June, 1974, pp. 132-146.

Lee, R. C. K., *Optimal Estimation, Identification, and Control*, MIT Press, Cambridge, MA, 1964.

Ljung, L., *System Identification: Theory for the User*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1987.

Ljung, L., and K. Glover, "Frequency domain versus time domain methods in system identification," *Automatica*, Vol. 17, 1981, pp. 71-86.

Matlab Users Guide, The MathWorks, Inc., Sherborn MA, 1987.

Mendel, J. M., *Discrete Techniques of Parameter Estimation*, Marcel Dekker, Inc., New York, 1973.

Mitchell, L. D., "Improved methods for the fast fourier transform (FFT) calculation of the frequency response function," *Journal of Mechanical Design*, Vol. 104, April 1982, pp. 277-279.

Moyar, G. J., W. D. Pilkey, and B. F. Pilkey, *Track/Train Dynamics and Design: Advanced Techniques*, Pergamon Press, Inc., New York, 1978.

Nguyen, V. V., and E. F. Wood, "Review and unification of linear identifiability concepts," *SIAM Review*, Vol. 24, No. 1, 1982.

Norton, J. P., "An investigation of the sources of nonuniqueness in deterministic identifiability," *Mathematical Biosciences*, Vol. 60, 1982a, pp. 89-108.

Norton, J. P., *An Introduction to Identification*, Academic Press, Orlando, FL, 1986.

Norton, J. P., "Letter to the editor," *Mathematical Biosciences*, Vol. 61, 1982b, pp. 295-298.

Norton, J. P., "Normal-mode identifiability analysis of linear compartmental systems in linear stages," *Mathematical Biosciences*, Vol. 50, 1980, pp. 95-115.

Paul, J. C., R. L. Gielow, A. E. Holmes, and P. K. Hickey, "Reduction of intermodal car aerodynamic drag through computerized flow simulation," *ASME Paper No. 83-RT-4*, New York, 1983.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: the Art of Scientific Computing*, Cambridge University Press, New York, 1986.

Rail Dynamics Laboratory Users Guide, RDL, TTC, Pueblo, CO, March, 1978.

Reid, J. G., "Identifiability and experimental design issues in linear system identification," *ASME Paper No. 79-WA/DSC-17*, New York, 1979.

Rost, B., and J. Leuridan, "A comparison of least-squares and total least-squares for multiple input estimation of frequency response functions," *ASME Paper No. 85-DET-105*, New York, 1985.

Sorensen, H. W., *Parameter Estimation*, Marcel Dekker, Inc., New York, 1980.

Trankle, T. L., "Practical aspects of system identification," *ASME Paper No. 79-WA/DSC-23*, New York, 1979.

Walter, E., *Identifiability of State Space Models,* Springer-Verlag, New York, 1982.

Walter, E., P. Bertrand, and G. LeCardinal, "Comments on 'Identifiability of linear and nonlinear dynamical systems'," *IEEE Trans. Autom. Control,* Vol. 24, No. 329-331, 1979.

Wilson, N. G., *Private Communication,* February, 1987.

Wilson, N. G., *Private Communication,* April, 1987.

Wormley, D. N., and P. A. Tombers, "Dynamic performance of single-axle freight trucks," *ASME Paper No. 83-RT-2,* New York, 1983.

Vita

Bradley M. Coffey was born to Dr. Joseph D. and Eloise G. Coffey on July 13, 1964 in Lima, Peru. His childhood was spent in several different locations in Peru, California, and Virginia. He graduated from Monacan High School in Richmond, Virginia in June 1982. He then attended Virginia Tech and received a Bachelor of Science degree *Magna Cum Laude* in Mechanical Engineering in July 1986. In September 1986, Bradley began his studies toward a Master of Science degree in Mechanical Engineering at Virginia Tech. Upon completion of his degree requirements, Bradley will enter the Civil Engineering department at Virginia Tech to pursue a doctorate in Environmental Engineering. After his doctorate, Bradley desires to teach and practice engineering in the Third World.