2011

Identifiability analysis of a tractor and single axle towed implement model

Simon Leroy Nielsen
Iowa State University

Identifiability analysis of a tractor and single axle towed implement model

by

Simon Leroy Nielsen

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Agricultural Engineering

Program of Study Committee:
Brian L. Steward, Major Professor
Stuart J. Birrell
James E. Bernard

Iowa State University

Ames, Iowa

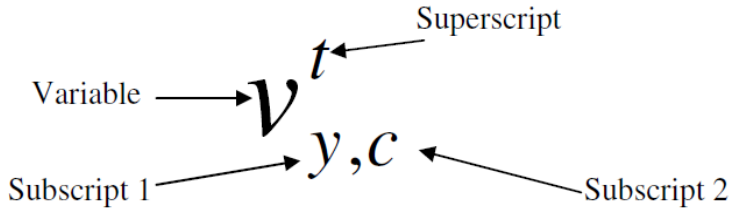
2011

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF VARIABLES



Variable: the variable itself

Superscript: denotes whether the variable is related to the tractor or the implement
t – tractor, i – implement

Subscript 1: specifies the coordinate axis the variable corresponds to
x – x-axis, y – y-axis, z – z-axis

Subscript 2: specifies the location the variable corresponds to
f – front, r – rear, c – center of gravity

Tractor-implement model

α	tire lateral slip angle
α_0	steady-state tire lateral slip angle
γ	yaw rate
δ	wheel steer angle
σ	tire relaxation length
ψ	heading angle
a	distance between front axle and CG of tractor
b	distance between rear axle and CG of tractor
c	distance between hitch point and CG of tractor
C_α	tire cornering stiffness
d	distance between hitch point and CG of implement
e	distance between rear axle and CG of implement
F	force
I	yaw moment of inertia
m	mass
u	longitudinal velocity
v	lateral velocity
x	position of a CG in the x-axis of the world coordinate system
y	position of a CG in the y-axis of the world coordinate system

Identification

$\boldsymbol{\theta}$	vector of model parameters to be estimated
θ_i	parameter of index i in θ

$\boldsymbol{\theta}^*$	true value of θ
$\widehat{\boldsymbol{\theta}}$	estimate of θ
σ	measurement error
χ_k^2	chi-square distribution with k degrees of freedom
CI	confidence interval bound
df	degrees of freedom
$\boldsymbol{e}_y$	output error
f_y	probability function
l	likelihood function
L	log likelihood function
$M(\boldsymbol{\theta})$	model with structure M and parameter vector θ
PL	profile likelihood abbreviation
$\mathbb{P}$	prior feasible set for θ
$\mathbb{R}^n$	set of real numbers in n-dimensional space
SNR	signal-to-noise ratio
t	time
$\boldsymbol{u}$	vector of controlled inputs to a model or system
$\mathbb{V}(\boldsymbol{\theta})$	neighborhood of parameters
$\boldsymbol{x}$	model state vector
$\boldsymbol{y}$	system output vector
$\boldsymbol{y_m}$	model output vector

ABSTRACT

The growing trend of model-based design in off-road vehicle engineering requires models that are sufficiently accurate for their intended application if they are to be used with confidence. Uncertain model parameters are often identified from measured data collected in experiments by using an optimization procedure, but it is important to understand the limitations of such a procedure and to have methods available for assessing the uniqueness and confidence of the results. The concept of model identifiability is used to determine whether system measurements contain enough information to estimate the model parameters. A numerical approach based on the profile likelihood of parameters was utilized to evaluate the local structural and practical identifiability of a tractor and single axle towed implement model with six uncertain tire force model parameters from tractor yaw rate and implement yaw rate data. The analysis first considered datasets generated from simulation of the model with known parameter values to examine the effect of measurement error, sampling rate, and input signal type on the identifiability. The results showed that the accuracy and confidence of identification tended to decrease as the quality, quantity, and richness of the data decreased, to the point that some of the parameters were considered practically unidentifiable from the information available. The profile likelihood plots also indicated potential opportunities for model reduction. Second, the analysis considered the identifiability of the model from two datasets collected during field experiments, and the results again indicated parameters that were practically unidentifiable from the information available. Overall, the study showed how different experimental factors can affect the amount of information available in a dataset for identification and that error in the measured data can propagate to error in model parameter estimates.

CHAPTER 1. GENERAL INTRODUCTION

1.1 Introduction

Off-road vehicle design and manufacturing companies are continually striving to meet customer needs by providing higher-quality and higher-performance products more quickly and at a lower cost. The term "off-road vehicles", used throughout this thesis, refers to the collection of ground vehicles used in the fields of e.g., agriculture, construction, logging, and mining (Wong, 2008). Companies in this industry are frequently faced with new engineering challenges, such as emissions regulations (ASABE, 2010), improvement of energy efficiency, increased safety requirements (Directive 2006/42/EC, 2006), and growing integration of electronic components for advanced control designs (Prabhu, 2007; DPNA, 2010).

The traditional engineering design process is iterative in nature, alternating between stages of design, testing, and analysis (Dieter and Schmidt, 2009). Intermediate forms of a final design, or prototypes, are often created and tested during the process to validate performance and obtain information for subsequent stages. However, design and analysis need not always be physical in nature. Engineers often develop mathematical models to characterize a physical system. In off-road vehicle applications, models have been developed for applications such as guidance controller design (Karkee and Steward, 2010), traction modeling (Book and Goering, 2000), ride and comfort evaluation (Ahmed and Goupillon, 1997), handling evaluation (Previati et al., 2007), and real-time driving simulators (Fales, et al., 2005; Hummel et al., 2005; Karimi and Mann, 2006; Karkee et al., 2009).

Advances in computer technology have had a major impact on engineering design and analysis over the last few decades (Dieter and Schmidt, 2009). An array of languages and software packages for modeling and simulation of systems have been developed to take advantage of faster and more-flexible computing platforms (Åström et al., 1998). These technologies have fostered the growing trend of model-based design strategies in the off-road vehicle industry (Prabhu, 2007). In general, a model-based approach utilizes characterizations of system behavior to meet specified design requirements (Wymore, 1993). Model-based design has the potential to reduce reliance on physical prototypes, which can lead to time and cost savings (Prabhu, 2007; Lennon, 2008). It also provides a means to explore potentially unsafe operating scenarios (Arikan, 2008; Lumpkin and Alford, 2010); for example, safety is a concern as tractors are engineered for higher-speed operation (Clay and Hemingway, 2001).

An extension of model-based design is virtual prototyping, in which design, analysis, and evaluation of products is performed in an immersive, interactive environment (Sastry and Boyd, 1998). As noted by Karkee (2009), "virtual prototyping (VP) is defined differently across disciplines and industries", but there are some elements that should be common to most dynamic system design and development applications, including: a modeling and simulation environment, a virtual reality environment, and a user interaction component. Virtual prototyping applications have been demonstrated in the off-road vehicle engineering context as well (Karkee et al., 2009).

Real-time simulation is a special class of simulation in that it enables real-time *analysis*. In many cases, real-time simulations of models are developed for hardware-in-the-loop studies, which incorporate physical hardware, controllers, and/or human operators. In

automotive applications, there are numerous examples of driving simulators. One of the most notable is the National Advanced Driving Simulator (NADS) at the University of Iowa, which provides a realistic environment for many types of studies (NADS, 2010). In off-road vehicle engineering specifically, there are several examples of the development of driving simulators (Fales, et al., 2005; Hummel et al., 2005; Karimi and Mann, 2006; Karkee et al., 2009). Real-time simulation faces the challenge of balancing model fidelity with computational resources, as the simulation time must keep up with the clock time. Specifically, for real-time simulation, the timeliness of an "answer" is as important as its correctness (Stankovich, 1988). This constraint can limit the fidelity of certain modeling domains in off-road vehicle systems, such as fluid power dynamics, tire-soil interaction, and 3D multibody dynamics. However, high-performance computing hardware, such as field-programmable gate arrays (FPGA), has recently been demonstrated for real-time simulation of a fluid power (hydraulic) system on an agricultural tractor (Karkee et al., 2010).

As modeling software capabilities and computer hardware capabilities have improved, there has been the potential for higher-fidelity and more-realistic models in engineering applications. Software capabilities often range from mathematical modeling to physical modeling to 3D computer-aided engineering (CAE) tools. Mathematical modeling pertains to the traditional characterization and derivation of mathematical system equations which are manually programmed for an application. Physical modeling is an object-oriented approach involving the development of model classes which represent physical components (Fritzson, 2004). It does not require the manual derivation of mathematical equations; instead, the software package interprets the representation and automatically forms the equations for simulation (Bernardin, 2009). Examples of physical modeling software include

Simscape (The MathWorks, Inc., Natick, MA) and the collection of platforms (e.g., Dymola, MapleSim, MathModelica, and OpenModelica) that use Modelica (The Modelica Association), an open physical modeling language. Lastly, CAE tools generally refer to the array of advanced software packages for performing, for example, finite element analysis (FEA), computational fluid dynamics (CFD) analysis, and multibody system (MBS) simulation, and they typically enable the integration of 3D CAD (computer-aided design) models and utilize elaborate graphical user interfaces; examples include ANSYS (ANSYS, Inc., Canonsburg, PA), LMS Virtual.Lab (LMS International, Leuven, Belgium), and MSC.Adams (MSC.Software Corporation, Santa Ana, CA). As engineers use the different modeling products provided by these software companies, the focus seems to be more on "representing" the system for simulation by numerical solvers than on obtaining closed-form mathematical expressions for analysis. In most cases, it is not possible to export a mathematical representation of a modeled system from the software package, so users must rely on the capabilities built into the software or apply numerical approaches.

An ongoing limitation in the advancement of model-based engineering design is the development of models with sufficient accuracy in which one can put confidence regarding their ability to characterize a system (Radhakrishnan and McAdams, 2005). Without sufficient confidence, the usefulness of a model is restricted, and there will be hesitancy to rely on it to drive decision-making in design. Validation processes can be run to ensure that a model can satisfactorily represent a physical system, at least within certain scenarios of interest (Ljung, 1999). However, this does not necessarily ease the initial process of model development.

1.2 Previous Work

Over the past few years, researchers in the Iowa State University Departments of Agricultural and Biosystems Engineering, Mechanical Engineering, and Electrical and Computer Engineering have worked together on several projects related to the modeling and simulation of off-road vehicle systems. Much of this research has been performed in collaboration with engineers at a leading company in the design and manufacture of off-road vehicles. A major outcome of this previous work was an architecture for modeling and real-time simulation of off-road vehicles in a virtual reality environment (Karkee et al., 2009).

However, as efforts to develop newer and higher-fidelity models have taken place, there have been many ongoing questions regarding how to improve and validate model accuracy for use in various applications. Karkee studied the modeling, identification, and analysis of a tractor and single axle towed implement system as part of his dissertation research (Karkee, 2009) at Iowa State University. After developing and analyzing three tractor-implement models of varying degrees of fidelity, it was found that a dynamic model with tire relaxation length dynamics included represented the system most accurately based on a comparison of frequency response. The parameters of the tire force model were found to be among those to which the vehicle response was most sensitive, but their values were also among those with the most uncertainty. An approach was developed and utilized to identify values for each of these tire model parameters based on system-level sensor data collected in field experiments. Overall, the parameter identification approach was shown to improve the ability for the model to represent system behavior compared to initially-selected parameter values. However, some of the parameters – particularly the tire relaxation length

parameters – were more difficult to estimate based on greater variation in their estimated values over several trials and larger standard deviation estimates.

Karkee's research demonstrated the potential of a parameter identification approach to address the challenge of developing more-accurate models for model-based design, but it also raised new questions regarding the practical limitations of such an approach in various applications. Although a parameter identification approach could always be used to obtain *some* set of more-or-less suitable values for a model's parameters, it was natural to question the accuracy, uniqueness, and overall confidence of those values as parameter estimates. For example, in later studies, it was observed that there was interaction between the different vehicle model parameters such that changes in one parameter value could be compensated, or "offset", by changes in other parameter values with little effect on the overall suitability of the model at representing system behavior. Also, further inspection of typical vehicle system sensor data from experiments revealed significant noise in some of the signals and called into question whether or not those signals contained enough information about the dynamics to be successfully used for identification.

In light of these observations, we desired to identify and demonstrate some type of methodology that could be used to assess the identification of parameters from experimental data; specifically, we desired to consider the effects of the model structure and experimental data properties on the identification of parameters. Such a methodology could have usefulness both before and after experimental data collection.

1.3 Research Objectives

The overall objective of the research in this thesis is to investigate methods that could be used to assess the process of identifying model parameters from experimental data. Based on those results, additional objectives are to investigate the identifiability of the tractor and single axle towed implement model's tire force model parameters from measured vehicle system data and to evaluate the effect of model structure as well as experimental factors such as measurement noise, data sampling rates, and input excitation type on the identification process. Overall, this thesis is intended to contribute additional knowledge to the process of parameter identification in the model-based design of off-road vehicle systems.

1.4 Organization of the Thesis

Chapter One provides a general introduction to model-based design, an overview of previous work, and the overall research objectives. Chapter Two presents a background on some of the methods used in the research, including some information on modeling, parameter identification, optimization, and identifiability. This chapter also includes a few simple examples that serve as case studies for the research that follows. Chapter Three is presented in the form of a paper intended for submission to a journal. That paper is focused on identifiability analysis of the tractor and single axle towed implement model mentioned above. Chapter Four presents some general conclusions derived from this research as well as recommendations for future work.

References

Ahmed, O. B., and J. F. Goupillon. (1997). Predicting the ride vibration of an agricultural tractor. *Journal of Terramechanics*, 34(1): 1-11.

Arikan, K. B. (2008). Identification of handling models for road vehicles. *PhD thesis*. Middle East Technical University.

ASABE. (2010). The EPA non-road diesel Tier 4 final rule. *Resource: Engineering & Technology for a Sustainable World*, 17(5): 10-13.

Bernardin, L. (2009). Speeding up the wheels of change. *Scientific Computing*, 26(1): 18-20.

Book, R. S., and C. E. Goering. (2000). A new traction model for crawler tractors. *Transactions of the ASABE*, 43(1): 39-46.

Clay R., and P. Hemingway. (2001). Engineering tractors for higher speeds. *ASAE Distinguished Lecture Series*. No. 25.

Dieter, G. E., and L. C. Schmidt. (2009). Engineering Design. Fourth Edition. New York: *McGraw-Hill*.

Directive 2006/42/EC. (2006). Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast). *Official Journal of the European Union*, L157/24-86.

DPNA. (2010). Furthering flexibility and safety. *Diesel Progress North America*, 76(3): 54-55.

Fales, R., E. Spencer, K. Chipperfield, F. Wagner, and A. Kelkar. (2005). Modeling and control of a wheel loader with human-in-the-loop assessment using virtual reality. *Journal of Dynamic Systems, Measurement, and Control*, 127: 415-423.

Fritzson, P. (2004). Principles of Object-Oriented Modeling and Simulation with Modelica 2.1. *Wiley-IEEE Press*.

Hummel, J., E. Baack, and J. Bernard. (2005). Operator in the loop simulation for farm and construction vehicles. *DSC 2005 North America – Orlando*, 421-432.

Karimi, D., and D. D. Mann. (2006). A study of tractor yaw dynamics for application in a tractor driving simulator. *Paper No. MBSK 06-113*, ASABE/CSBE North Central Intersectional Meeting.

Karkee, M. (2009). Modeling, identification and analysis of tractor and single axle towed implement system. *PhD dissertation*, Iowa State University.

Karkee, M., B. L. Steward, A. G. Kelkar, and Z. T. Kemp II. (2009). Modeling and real-time simulation architectures for virtual prototyping of off-road vehicles. *Virtual Reality*, 15: 83-96.

Karkee, M., and B. L. Steward. (2010). Study of the open and closed loop characteristics of a tractor and a single axle towed implement system. *Journal of Terramechanics*, 47(6): 379-393.

Karkee, M., M. Monga, B. L. Steward, J. Zambreno, and A. G. Kelkar. (2010). Real-time simulation and visualization architecture with field programmable gate array (FPGA) simulator. *Proceedings of ASME 2010 World Conference on Innovative Virtual Reality*, May 12-14, Ames, IA.

Lennon, T. (2008). Model-based design for mechatronic systems. *EngineerIT*, April: 16-20.

Ljung, L. (1999). System Identification: Theory for the User. Second Edition. Englewood Cliffs, NJ: *Prentice Hall*.

Lumpkin, E., and C. Alford. (2010). How to make virtual prototyping better than designing with hardware: Part 2. *EE Times*. Available at: http://www.eetimes.com/design/embedded/4200574/How-to-make-virtual-prototyping-better-than-designing-with-hardware-Part-2. Accessed: 9 June, 2011.

NADS. (2010). National Advanced Driving Simulator Overview. *The National Advanced Driving Simulator at The University of Iowa*. Available at: http://www.nads-sc.uiowa.edu/press/pdf%5CNADS_Overview.pdf. Accessed: 2 June, 2011.

Prabhu, S. (2007). Model-based design for off-highway machine systems development. *SAE 2007 Commercial Vehicle Engineering Congress & Exhibition.* SAE Paper No. 2007-01-4248.

Previati, G., M. Gobbi, and G. Mastinu. (2007). Farm tractor models for research and development purposes. *Vehicle System Dynamics*, 45: 37-60.

Radhakrishnan, R., and D. A. McAdams. (2005). A methodology for model selection in engineering design. *ASME Journal of Mechanical Design*, 127: 378-387.

Sastry, L., and D. R. S. Boyd. (1998). Virtual environments for engineering applications. *Virtual Reality*. 3: 235-244.

Stankovic, J. (1988). Misconceptions about real-time computing: a serious problem for next generation systems. *IEEE Computer*, 21(10): 10-19.

Wong, J. Y. (2008). Theory of Ground Vehicles. Fourth Edition. Hoboken, NJ: *John Wiley & Sons, Inc.*

Wymore, A. W. (1993). Model-Based Systems Engineering. Boca Raton, FL: *CRC Press.*

Åström, K. J., H. Elmqvist, and S. E. Mattsson. (1998). Evolution of continuous-time modeling and simulation. *In: Proceedings of the 12th European Simulation Multiconference, ESM'98.* June 16-19, Manchester, UK.

CHAPTER 2. BACKGROUND

The work presented in this thesis takes a closer look at some issues related to the identification of a parametric vehicle model from output data. It is not necessarily focused on the development or validation of off-road vehicle or tire models but is based upon recent work done in this area. It examines the topic of model identifiability in the context of a tractor and single axle towed implement identification experiment in order to determine whether there is a unique solution for the identified parameters. This chapter presents background information on some methods in the literature that are used later in this thesis.

2.1 Parameter Identification

As described by Walter and Pronzato (1997), physical systems are generally modeled in continuous time and described by a set of differential equations,

$$\dot{\boldsymbol{x}}(t) = f\big(\boldsymbol{x}(t), \boldsymbol{\theta}, \boldsymbol{u}(t)\big) \tag{2.1}$$

$$\boldsymbol{y}_m(t) = h\big(\boldsymbol{x}(t), \boldsymbol{\theta}, \boldsymbol{u}(t)\big) \tag{2.2}$$

where $\boldsymbol{x}$ is the state vector, $\boldsymbol{\theta}$ is the parameter vector, $\boldsymbol{u}$ is the vector of controlled inputs, t is time, and $\boldsymbol{y}_m$ is the vector of model outputs.

At a high level, several different model types can be considered, each having advantages and disadvantages in different applications (Walter and Pronzato, 1997; Ljung, 1999; Bohlin, 2006). The most common type is the "white box" model, which is guided by first principles such as conservation and balance to represent system phenomena. Any model parameters have physical meaning and can be measured and specified directly. However, it is difficult to completely describe many complex systems in this manner. At the other end of the modeling spectrum is the "black box" model, in which an arbitrary mathematical

structure is used to fit an input to an output recorded in experimental data. The model parameters generally have no physical interpretation, so their values lend less direct insight into the underlying system. From an engineering design standpoint, it is often more beneficial to examine the impact of a physical parameter change on performance. However, black box approaches can be very effective at modeling observed system behavior. In between white box and black box models is the "grey box" modeling approach. A first principles model structure is used to explain most, if not all, of the system behavior. Some of the model parameter values may be known with greater certainty, but other parameter values may be unknown. The unknown parameters are then identified, or "estimated", from experimental data. The main advantage of a grey-box model is that its parameters retain physical meaning, yet it has been calibrated to match observed system behavior.

A typical way to obtain the parameters for a given model structure is to find the set of values for which the model output most closely represents the actual system output for a given input (Walter and Pronzato, 1997). This general process is illustrated in Figure 2.1.

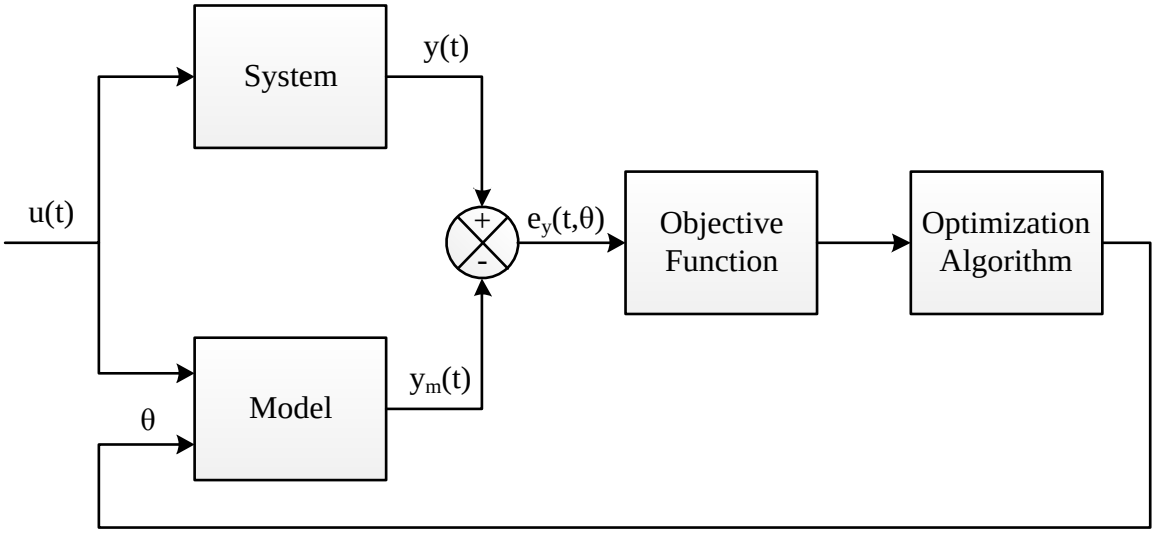


Figure 2.1 – The general parameter identification process. (Adapted from similar figures given by Walter and Pronzato (1997).)

Closeness of representation may be determined by comparing the time history of one or more sensed outputs of the system with the time history of the same outputs of the model. For a common input vector, $\boldsymbol{u}$, the error vector, $\boldsymbol{e}_y$, between the system output vector, $\boldsymbol{y}$, and the corresponding model output vector, $\boldsymbol{y}_m$, is calculated as,

$$\boldsymbol{e}_y(t, \boldsymbol{\theta}) = \boldsymbol{y}(t) - \boldsymbol{y}_m(t, \boldsymbol{\theta}) \tag{2.3}$$

In an effort to obtain the best estimate of parameter values, $\widehat{\boldsymbol{\theta}}$, for the model to characterize the system, an objective function will be formulated that calculates a scalar value as a function (e.g., the sum of squares) of the output error, $\boldsymbol{e}_y$. The purpose of the objective function is to quantify the suitability of the model with a particular set of parameter values, and the calculation and weighting of error can be performed in any way that emphasizes the "objective" of the optimization. Therefore, obtaining the optimal set of model parameters (i.e., the one with the least error) becomes a problem of minimizing the objective function. Since an exhaustive search of the parameter space is rarely practical, an optimization algorithm will generally be used to search the parameter space for the minimum value of the objective function. Such a minimization problem can be approached with one of many algorithms available in the literature which search the parameter space for the optimal set of parameter values.

2.2 Optimization Algorithms

As mentioned above, identifying the set of parameter values for which a model best represents a physical system can be approached as an optimization problem with the objective of minimizing the error between the outputs of the model and system. For a problem with one or two parameter values to optimize, the parameter space can be

envisioned as a one- or two-dimensional "landscape", respectively, with an additional, vertical dimension reflecting the value of the objective function at any unique parameter value or pair of values. The topography of the landscape will generally include regions of relatively lower objective function values, or minima, that optimize, either locally or globally, the parameters for that particular objective function. The general goal of optimization is to locate the global minimum of the landscape formed by that objective function. For parameter sets of dimension greater than two, the previously-mentioned landscape is less intuitive to visualize, but the goal of optimization within that parameter space is the same.

Various optimization methods have been suggested as being more or less effective for different problem types, parameter set dimensions, and optimization scopes (Venkataraman, 2009). Among the different optimization methods available, two classes emerge: local techniques and global techniques.

2.2.1 Local Optimization

Local optimization techniques are usually deterministic in nature. They are, in general, not likely to converge to the global optimum, and their results are highly dependent upon the initial values, or "starting point", of the design variables. Repeated application of a local optimization technique from different initial values in the parameter space will tend to increase the chances of reaching a global optimum; however, this outcome is not guaranteed, and this approach becomes more difficult to implement as the number of parameters increases. Common techniques include Newton's method and gradient descent (Nocedal and Wright, 1999; Venkataraman, 2009).

2.2.2 Global Optimization

Global optimization techniques are usually stochastic in nature. They search the parameter space in a heuristic manner and are more likely to reach a global optimum, although this is also not guaranteed. The main drawback to global techniques is computational expense associated with the large number of function calls typically required during optimization. Common examples include evolutionary computation and simulated annealing (Venkataraman, 2009).

The genetic algorithm (GA) approach (Goldberg, 1989), a subset of the broader class of evolutionary computation, was inspired by the natural genetic processes of inheritance, selection, crossover, and mutation, and has shown considerable usefulness in global optimization and search applications since the 1980s, aided by the widespread advance of computer technology (Goldberg, 1994). A number of applications in vehicle dynamics have made use of GA for optimization problems. It has been used for optimization of vehicle trajectory simulations (Bernard et al., 1998; Fittanto and Puig-Suari, 2000; Bernard and Balling, 2004) as well as automotive design applications (Fujita et al., 1998; Hoffmeister and Bernard, 1998).

Examples of the use of GA specifically for vehicle model parameter identification based on experimental data are available in the literature; although, as noted by Arikan (2008), it seems to be less common than the use of other identification techniques. A GA approach has been used for the identification of two and three degree of freedom handling models for road vehicles based on experimental data (Arikan, 2008) as well as simulated data obtained from high-fidelity multibody models (Bolhasani and Azadi, 2002). A GA approach has also been used for identification of a drivetrain model (Maclay and Dorey, 1993),

identification of a nonlinear vehicle ride model based on simulated data (Alasty and Ramezani, 2002), and identification of parameters for a semi-empirical tire model based on the results of finite element analysis (FEA) of a tire in sand (Grujicic et al., 2010). In off-road vehicle engineering applications, there is even less evidence of the use of GA for optimization problems. A GA approach was shown to improve the identification of hydraulic system parameters from experimental data compared to a manual search (Book, 1996). A GA approach was also used to identify the parameters for a traction model of a crawler tractor based on data collected in field experiments (Book and Goering, 2000).

In the trend toward higher-fidelity and more-realistic models for use in design in off-road vehicle applications, GA may be a useful technique for parameter identification that has not been fully considered. GA optimization is non-deterministic, and the algorithm's direct interaction with a model is as straightforward as proposing a set of parameters and evaluating the simulation results (Goldberg, 1994). This nature of interaction means that global optimization approaches can be applied to almost any model type without needing to evaluate or approximate gradients or Hessians of the objective function. A GA approach was used in Chapter 3 of this thesis to obtain an initial parameter set for a tractor-implement model with respect to experimental data.

2.3 Identifiability

Although the gray-box approach to system modeling may promote model accuracy for use in model-based design, success is not necessarily guaranteed for identification of any arbitrary model's parameters from any experimental dataset. Model identifiability analysis is used to determine whether system measurements contain enough information to estimate the

model parameters. That is, identifiability refers to the uniqueness of a parameter vector $\widehat{\boldsymbol{\theta}}$ as an estimate of the true parameter vector $\boldsymbol{\theta}^*$ in a model M representing a physical system (Walter and Pronzato, 1997). Within this field, there are two subtypes frequently referred to as structural identifiability and practical identifiability in the literature.

2.3.1 Analytical Approaches

2.3.1.1 Structural Identifiability

From a structural standpoint, a model may be parameterized in a way that one or more parameters cannot be uniquely determined from the output. Furthermore, often only a limited number of the model states can be measured on an actual system. Structural identifiability analysis is conducted independent of any parameter values or experimental data and is concerned with determining the ability to identify a model in ideal conditions – that is, with no error in modeling the system, no noise in the data, and with input and output measurement times "chosen at will" (Walter and Pronzato, 1997). Therefore, it is also referred to as the theoretical or *a priori* identifiability. It is considered to be the "qualitative" aspect of experimental design for parameter identification (Walter and Pronzato, 1990).

Formal definitions for structural identifiability are given by e.g., Walter and Pronzato (1997) and Ljung (1999). The following explanation follows closely from the definitions given by Walter and Pronzato (1997). If the condition

$$M\left(\widehat{\boldsymbol{\theta}}\right) = M(\boldsymbol{\theta}^*) \Rightarrow \hat{\theta}_i = \theta_i^* \tag{2.4}$$

holds for almost any $\boldsymbol{\theta}^*$ in $\mathbb{P}$ (the prior feasible set for $\boldsymbol{\theta}$; $\mathbb{P} = \mathbb{R}^n$ unless otherwise stated), then a parameter θ_i is classified as *structurally globally identifiable*. In other words, "identical input-output behavior" of two identical model structures implies that the estimated parameter set $\widehat{\boldsymbol{\theta}}$ is unique and corresponds to the true parameter set $\boldsymbol{\theta}^*$. Furthermore,

structural global identifiability of each parameter θ_i in $\boldsymbol{\theta}$ is a necessary condition for structural global identifiability of the model structure. The condition "almost any $\boldsymbol{\theta}^*$" functions to exclude atypical parameter values that may cause other parameters to become unidentifiable. If a model structure cannot be classified as globally identifiable, it may be possible to verify the model's local identifiability for some neighborhood $\mathbb{V}(\boldsymbol{\theta}^*)$ around the true parameter set. If Eq. (2.4) holds for $\widehat{\boldsymbol{\theta}} \in \mathbb{V}(\boldsymbol{\theta}^*)$ then a parameter θ_i is classified as *structurally locally identifiable*. Each parameter θ_i in $\boldsymbol{\theta}$ must be at least structurally locally identifiable for the model structure to be classified as structurally locally identifiable. Consequently, local identifiability is a necessary condition for global identifiability. If there does not exist a neighborhood $\mathbb{V}(\boldsymbol{\theta}^*)$ for which Eq. (2.4) holds, then a parameter θ_i is classified as *structurally unidentifiable*. A model structure is structurally unidentifiable if one or more of its parameters is unidentifiable.

A number of analytical methods have been developed for investigating structural identifiability and are demonstrated for various applications in the literature. For linear models, two common methods for testing structural identifiability include the Laplace transform, or "transfer function", approach (Bellman and Åström, 1970) and the similarity transformation approach. An approach for local identifiability of linear models in state-space format is to assemble the Markov parameter matrix and determine if there is a one-to-one mapping from the parameter space to the Markov parameters; this is ensured if the Jacobian of the matrix is full rank (Grewal and Glover, 1976). For nonlinear models (and linear models), the Taylor series approach, local state isomorphism approach, and elimination theory (e.g., differential algebra) are available. More information on these methods is given by Walter and Pronzato (1997).

Consider, for example, the identification of a simple mechanical system shown in Figure 2.2 consisting of a mass, spring, and damper with one translational degree of freedom.

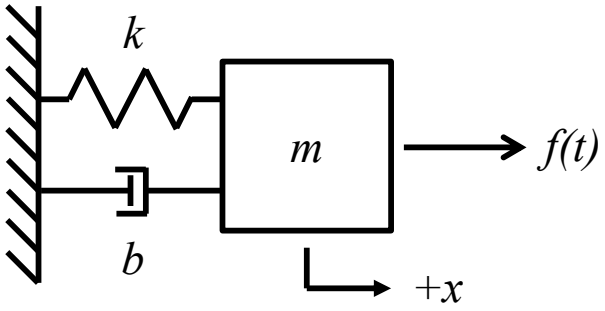


Figure 2.2 – A simple mechanical system with mass *m*, spring constant *k*, and damping coefficient *b*. The system has a translational degree of freedom along the *x*-axis and is acted upon by a horizontal force $f(\mathbf{t})$.

This system can be modeled mathematically by constructing a free body diagram for the mass and summing forces in the *x*-direction, which leads to the following second-order model relating the force *f(t)* to the displacement, *x*, of the mass from its equilibrium position,

$$m\ddot{x} + b\dot{x} + kx = f(t) \tag{2.5}$$

In this case, the objective is to identify the values of all three parameters, *m*, *k*, and *b*, by measuring the displacement, *x*, with a sensor as the force, *f(t)*, is input to the system. One may seek to determine if there is a globally unique solution for these three parameters based on the input and output used for this experiment. Since the model is a linear, time-invariant, single-input, single-output structure, the Laplace transform (or "transfer function") approach originally proposed by Bellman and Åström (1970) can be used. Assuming initial conditions of zero for the position and velocity states of the mass and taking the Laplace transform of Eq. (2.5), the open-loop transfer function for this system in canonical form is,

$$G(s) = \frac{X(s)}{F(s)} = \frac{\frac{1}{m}}{s^2 + \frac{b}{m}s + \frac{k}{m}} \tag{2.6}$$

The structural, global identifiability requirement, $M(\hat{\boldsymbol{\theta}}) = M(\boldsymbol{\theta}^*)$, is equivalent to,

$$\frac{1}{\hat{m}} = \frac{1}{m^*} \tag{2.7}$$

$$\frac{\hat{b}}{\hat{m}} = \frac{b^*}{m^*} \tag{2.8}$$

$$\frac{\hat{k}}{\hat{m}} = \frac{k^*}{m^*} \tag{2.9}$$

which has a unique solution,

$$\hat{\boldsymbol{\theta}} = (m^*, k^*, b^*) \tag{2.10}$$

All three parameters of the model and, therefore, the model itself are structurally globally identifiable. Therefore, the true values for *m*, *k*, and *b*, can be determined uniquely from noise-free input and output data.

However, consider a similar mechanical system shown in Figure 2.3 which has two springs in parallel between the fixed ground and the mass.

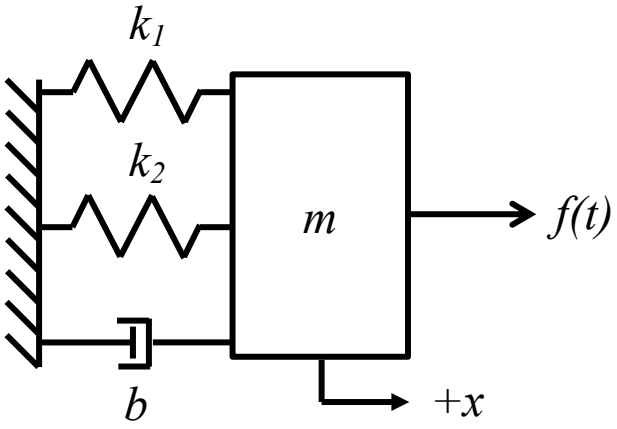


Figure 2.3 - A simple mechanical system with mass m, spring constants k_1 and k_2, and damping coefficient b. The system has a translational degree of freedom along the x-axis and is acted upon by a horizontal force $f(t)$.

The second-order mathematical model for this system is,

$$m\ddot{x} + b\dot{x} + (k_1 + k_2)x = f(t) \tag{2.11}$$

and the transfer function for this system is,

$$G(s) = \frac{X(s)}{F(s)} = \frac{\frac{1}{m}}{s^2 + \frac{b}{m}s + \frac{k_1+k_2}{m}} \tag{2.12}$$

The structural, global identifiability requirement, $M(\widehat{\boldsymbol{\theta}}) = M(\boldsymbol{\theta}^*)$, is equivalent to,

$$\frac{1}{\hat{m}} = \frac{1}{m^*} \tag{2.13}$$

$$\frac{\hat{b}}{\hat{m}} = \frac{b^*}{m^*} \tag{2.14}$$

$$\frac{\hat{k}_1 + \hat{k}_2}{\hat{m}} = \frac{k_1^* + k_2^*}{m^*} \tag{2.15}$$

The mass and damping coefficient are both structurally identifiable, but each spring constant estimate could take on one of many feasible values as long as $\hat{k}_1 + \hat{k}_2 = k_1^* + k_2^*$. The relationship between the input and output only depends on the effective spring constant of the system. It would be necessary to have additional information about k_1 or k_2 (e.g., from a previous experiment) in order for the model to be identifiable. In this case, it is relatively straightforward to recognize, based on visual inspection and experience, that it would be impossible to uniquely identify spring constants for the two springs in parallel. However, the ability to make this judgment for arbitrary models may be limited as model complexity increases. Note that sensitivity analysis alone cannot be used to determine structural identifiability; the output, x, has a nonzero sensitivity to k_1 and k_2 in both transient and static response to an input force $f(t)$, but this is not sufficient for unique identification of these two parameters.

2.3.1.2 Practical Identifiability

In parameter identification, there are concerns regarding the amount of information contained in the actual experimental data (Ljung, 1999). Even if a model has been determined to be structurally identifiable, this does not necessarily guarantee successful estimation from measured data. The ideal conditions in Section 2.3.1.1 under which structural identifiability is evaluated are certainly not characteristic of actual parameter identification experiments. Practical identifiability, however, considers model identifiability in light of the characteristics of the experimental data used for identification (Balsa-Canto and Banga, 2010). This is also referred to as the *a posteriori* identifiability or the "quantitative" aspect of experimental design for parameter identification and is also related to a field known as optimal experimental design (Walter and Pronzato, 1990). Experimental data properties considered in identification experiments often include quality, quantity, and richness. Data quality refers to the presence of error in the output data. Quantity refers to the actual number of data points available. Richness of data is related to the manner in which a system input is excited; richer datasets are generated by inputs that contain spectral content across the bandwidth of the model and persistently excite the system (Ljung, 1999). These data properties as well as the interaction between parameters can affect the certainty of parameter estimates. Therefore, it is possible for a structurally identifiable parameter to be considered practically unidentifiable once experimental data is introduced. However, unlike structural identifiability, practical identifiability is not as clear of a "yes-or-no" question. As noted by Raue et al. (2009), the literature does not seem to provide as clear of criteria defining practical identifiability versus unidentifiability.

Sensitivity analysis is related to practical identifiability analysis, but it does not provide quite the same type of information. Sensitivity analysis is concerned with variation in the system output due to variation in system parameters and determines the influence of parameters on system behavior (Banks, 1998). There are many practical applications of sensitivity analysis (Saltelli et al., 2000). In product engineering, sensitivity analysis can aid in determining which aspects of a system must be designed and manufactured with the highest precision. In dynamic system modeling and simulation, sensitivity analysis can aid in determining which aspects of a system must be investigated and measured with greatest certainty (Karkee and Steward, 2010). However, practical identifiability analysis is concerned with uncertainty in the identification of system *parameters* due to variation in the measured system *output* and takes into account the interaction between parameters.

The statistical aspects of parameter identification from experimental data have been considered with maximum likelihood principles (Ljung, 1999), for which a number of inferential methods are available (Meeker and Escobar, 1998). The following explanation of maximum likelihood estimation is based on descriptions given by Vardeman and Jobe (2001). For a set of observed data $\boldsymbol{y}$ (with each observation assumed independent and identically distributed) and a model with parameter vector $\boldsymbol{\theta}$, the probability function $f_y(\boldsymbol{\theta}; \boldsymbol{y})$ of a model taking the value $\boldsymbol{y}$ for various parameter vector values is called the likelihood function, and its natural logarithm is called the log likelihood function,

$$L(\boldsymbol{\theta}) = \ln\left(f_y(\boldsymbol{\theta}; \boldsymbol{y})\right) \tag{2.16}$$

Therefore, the goal is to determine the set of parameters $\boldsymbol{\theta}$ that maximizes L. The maximum likelihood estimate $\widehat{\boldsymbol{\theta}}$ is the one that maximizes the probability of the observed data (Ljung, 1999).

For large samples, likelihood-based confidence intervals can be defined to contain a region of parameter values around the maximum likelihood estimate $\widehat{\boldsymbol{\theta}}$ of dimension k for which the likelihood is highest (Vardeman and Jobe, 2001). In other words, the confidence region is defined as,

$$\left\{\boldsymbol{\theta}\middle|L(\boldsymbol{\theta}) > L\left(\widehat{\boldsymbol{\theta}}\right) - \Delta L\right\} \tag{2.17}$$

where ΔL is a value appropriate for the confidence level desired. This concept is illustrated in Figure 2.4 for a single parameter θ.

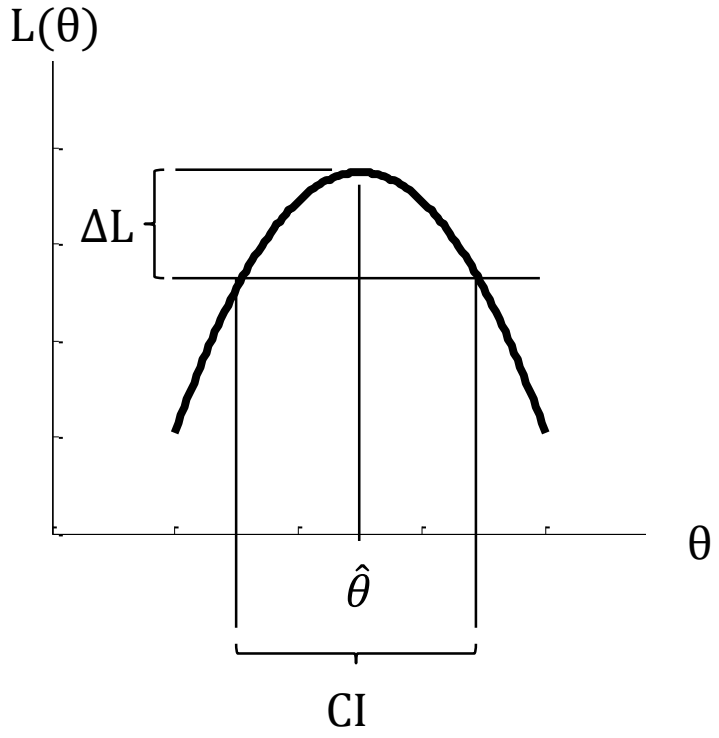


Figure 2.4 - Plot of a log likelihood function $L(\boldsymbol{\theta})$ (the parabolic curve) and indication of a likelihood-based confidence interval (CI) for $\boldsymbol{\theta}$ around its maximum likelihood estimate $\widehat{\boldsymbol{\theta}}$ based on a value ΔL.

An approximate $100(1-\alpha)\%$ likelihood-based confidence region is obtained when

$$\Delta L = \frac{1}{2}U \tag{2.18}$$

where U is the $(1-\alpha)$ quantile of the χ_k^2 distribution.

For two parameters θ_1 and θ_2, two-dimensional likelihood-based confidence regions are visualized around the maximum likelihood estimate. The log likelihood value is plotted on the axis coming out of the page, and confidence regions are bounded by contours of constant value $L(\boldsymbol{\theta})$ as shown in Figure 2.5. This theory extends similarly to parameters sets of dimension greater than two, but visualization of the confidence region is not as intuitive.

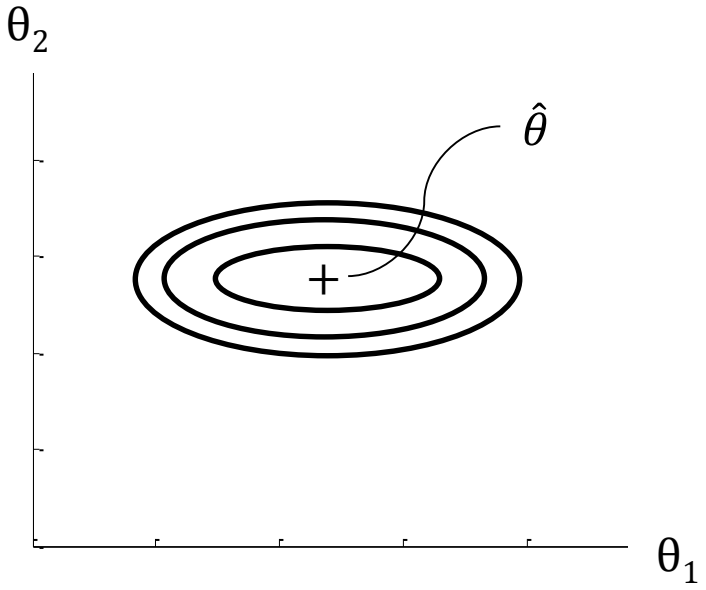


Figure 2.5 – Plot of likelihood-based confidence regions for two parameters. $\boldsymbol{L(\theta)}$ is plotted on the axis coming out of the page. Each parameter confidence region is bounded by a contour of constant value $\boldsymbol{L(\theta) = L(\hat{\theta}) - \Delta L}$

The Fisher Information Matrix (FIM) is a measure used for testing practical identifiability (Balsa-Canto and Banga, 2010) and determining an optimal experimental design for identification (Walter and Pronzato, 1997). As described by Meeker and Escobar (1998), the FIM is calculated as an expectation of information in future data,

$$FIM = E\left[-\frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\, \partial \boldsymbol{\theta}^T}\right] \tag{2.19}$$

The precision of estimation is implied by the curvature in the likelihood. Greater precision is indicated by larger second-derivatives of the log likelihood $L(\boldsymbol{\theta})$. If data is available, the local FIM can be computed by evaluating Eq. (2.19) at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. The inverse of the FIM provides an approximate large-sample covariance matrix.

2.3.2 Numerical Approaches

Application of analytical identifiability approaches to large, complex models can be mathematically impractical, even with the help of symbolic math software. Numerical approaches for local, structural identifiability have been proposed in mechanical system contexts (Serban and Freeman, 2001; Jiafan et al., 2010).

Identifiability of dynamic models is an active topic of research in the field of systems biology. Researchers in this field develop mathematical models of biological reaction networks and identify model parameters based on experimental observations. According to Raue et al. (2009), their reaction networks permit only a limited number of outputs to be measured, and experimental data is often of insufficient quantity and quality for parameter identification; furthermore, the size and complexity of their mathematical models often renders analytical identifiability methods inappropriate.

Therefore, numerical approaches for detecting unidentifiability of models have been investigated (Hengl et al., 2007; Raue et al., 2009). To address the problems stated above, Raue et al. (2009) proposed an approach to evaluate both structural and practical identifiability of model parameters based on their profile likelihood. The approach is data-based, enabling practical identifiability to be used to evaluate experimental factors such as

data quantity, noise, and system input. An important feature of the profile likelihood identifiability approach is that it can be applied to arbitrary model types. At a minimum, it is only necessary to be able to call the model with a specific parameter set and to obtain the outputs for objective function calculation. This particular aspect can be favorable for any method used in model-based design. Since modeling is done in various software formats, it is not always possible to obtain a closed-form, mathematical representation for a system.

Based on a review of the literature pertaining to identifiability analysis and the feasibility of different methods in the context of off-road vehicle modeling applications, the profile likelihood approach proposed by Raue et al. (2009) was utilized for the identifiability analysis in this thesis. In the following, this approach is explained and applied to some simple mechanical system examples.

2.3.2.1 Profile Likelihood Approach

Raue et al. (2009) and (2011) described a numerical approach to local structural and practical identifiability based on the profile likelihood of the model parameters. A detailed description of the approach can be found in those studies but is summarized as follows. For the optimization problem, they considered an objective function which is the weighted sum of squared residuals

$$\chi^2(\boldsymbol{\theta}) = \sum_{k=1}^{m}\sum_{l=1}^{d}\left(\frac{y_{kl}^D - y_k(\boldsymbol{\theta}, t_l)}{\sigma_{kl}^D}\right)^2 \tag{2.20}$$

where k is the index of m outputs measured, l is the index of d data points collected, y_{kl}^D is an experimental data point, y_k is a model output, and σ_{kl}^D is the corresponding measurement error of a data point. Assuming that the noise on the measurements is normally distributed, $\epsilon \sim N(0, \sigma^2)$, minimization of this objective function yields maximum likelihood estimates of

the parameter set, $\boldsymbol{\theta}$. Although asymptotic confidence intervals for the parameters can be obtained based on a quadratic approximation of the likelihood at the estimated parameter values if the model "sufficiently describes the experimental data", Raue et al. (2009) acknowledged that this approximation may not hold as well for cases with data of lower quality and/or quantity. For those cases, confidence intervals based on a "threshold" in the likelihood were recommended, defined by

$$\left\{\boldsymbol{\theta} | \chi^2(\boldsymbol{\theta}) - \chi^2\left(\hat{\boldsymbol{\theta}}\right) < \Delta_\alpha\right\} \tag{2.21}$$

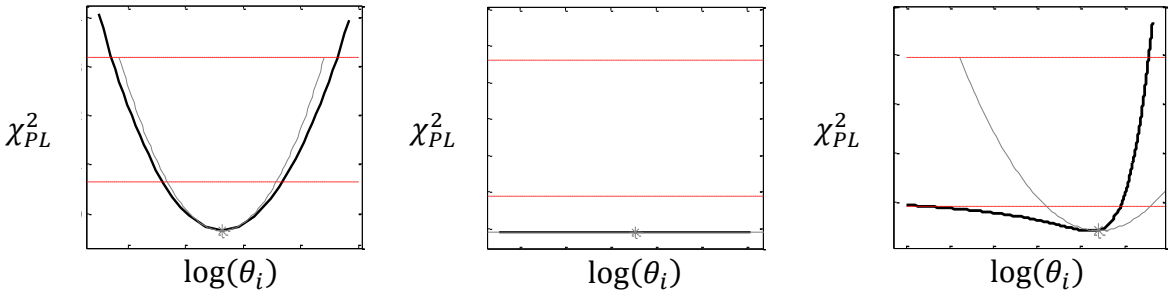
$$\Delta_\alpha = Q\left(\chi^2_{df}, 1-\alpha\right) \tag{2.22}$$

where Δ_α is the $1-\alpha$ quantile of the χ^2-distribution with df degrees of freedom.

Raue et al. (2009) sought to efficiently search the parameter space around each parameter estimate by "exploring the parameter space for each parameter in the direction of least increase in χ^2". The profile likelihood was selected for that objective. This computation individually increments each parameter in increasing and decreasing directions around its estimate, reoptimizing all of the other parameters to the data and recording the χ^2 (objective function) value at each step. Therefore, the approach is able to capture the effects of parameter sensitivity as well as parameter correlation on the identification of model parameters. The computation produces a profile likelihood plot for each parameter, showing how its likelihood changes with respect to the parameter values. Based on Eqs. (2.21) and (2.22), upper and lower confidence bounds for a parameter are determined by the locations at which the likelihood crosses a certain χ^2 threshold.

Assessment of parameters' local identifiability can then be made from their likelihood-based confidence intervals in logarithmic space. Raue et al. (2009) defined a

parameter as identifiable if it has finite confidence bounds, i.e., a profile likelihood that reaches a specific χ^2 threshold (upper red dashed line) as shown in Figure 2.6a.



a) b) c)

Figure 2.6 – Example profile likelihood plots (in black) for a) an identifiable parameter, b) a structurally unidentifiable parameter, and c) a practically unidentifiable parameter. The gray curve results from a quadratic approximation of the likelihood.

A completely flat profile likelihood with no minimum, as shown in Figure 2.6b, is structurally unidentifiable, indicating a functional relation between parameters such that a change in one parameter value can be compensated by a change in at least one other parameter with no increase in the objective function. Acknowledging that (compared to structural identifiability) the literature does not provide as clear of a definition for practical unidentifiability, Raue et al. (2009) defined a parameter as practically unidentifiable if it has an infinite upper and/or lower confidence bound but the likelihood has a definite minimum value. An example of this case is shown in Figure 2.6c. Even if a parameter is deemed identifiable, the precision of its estimation can be assessed by the width of its confidence interval.

This profile likelihood approach described above is implemented into the third-party PottersWheel mathematical modeling toolbox (Maiwald and Timmer, 2008) for MATLAB (The MathWorks, Inc., Natick, MA). Although the toolbox is tailored specifically toward the

systems biology community, it has the capability to handle general mathematical models defined as a set of ordinary differential equations as well. In addition, the toolbox has many other functionalities that are useful in mathematical modeling, parameter identification, and model analysis.

2.3.2.2 Examples

As an example of the profile likelihood identifiability approach implemented in PottersWheel, consider again the structurally unidentifiable mass-spring-damper system in Figure 2.3, modeled by the second-order equation,

$$m\ddot{x} + b\dot{x} + (k_1 + k_2)x = f(t) \tag{2.23}$$

The PottersWheel toolbox requires mathematical models to be entered in a specific M-file format compatible with its functions; particularly, the model must be entered as a set of ordinary differential equations. Reduction of order for Eq. (2.23) results in,

$$\dot{y}_1 = y_2 \tag{2.24}$$

$$\dot{y}_2 = \frac{f}{m} - \frac{b}{m}y_2 - \frac{(k_1 + k_2)}{m}y_1 \tag{2.25}$$

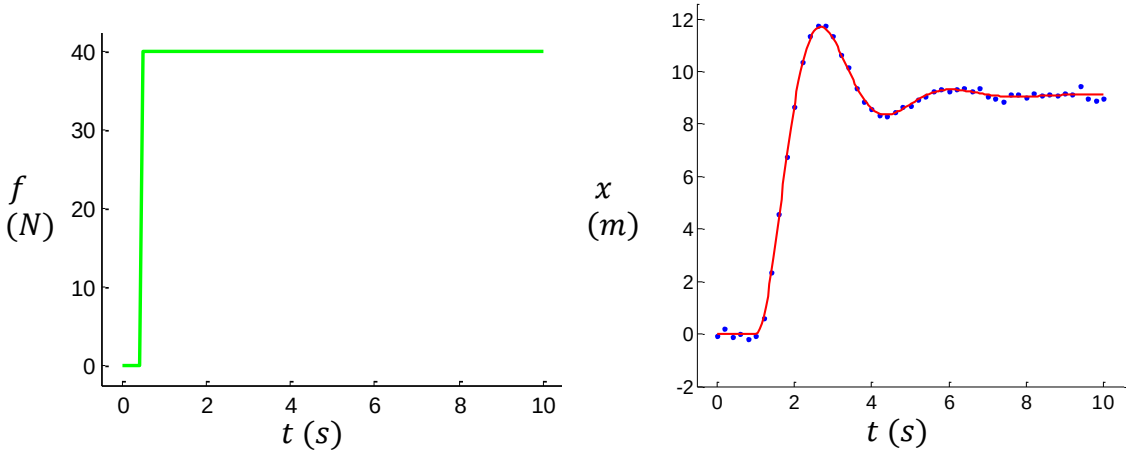
where $y_1 = x$ and $y_2 = \dot{x}$. The model input, $f(t)$, was defined using a driving input function with predefined input types. All four model parameters were considered to be unknown (or "free"), but within the bounds given in Table 2.1.

Table 2.1 - Parameter bounds for the simple mass-spring-damper system example.

Parameter	Units	Lower Bound	Upper Bound
m	kg	0.01	5.00
k_1	N/m	0.01	5.00
k_2	N/m	0.01	5.00
b	N-s/m	0.01	5.00

Nominal values of $m = 1.1$, $k_1 = 3.0$, $k_2 = 1.4$, and $b = 1.6$ were specified. The model output, x, was defined with an error model and sampled at 5 Hz. The error model was Gaussian and had a standard deviation of approximately 0.1 m (compared to a change in the steady-state position of the mass of approximately 9 m) in order to produce a dataset with low noise amplitude, i.e., a high signal-to-noise ratio.

After loading the model into the PottersWheel graphical user interface (GUI), the option to create simulated data was used. A step input force of 40 N was applied. The simulation used the nominal values of the four free parameters and applied the specified error model to the output. The built-in CVODES solver for ordinary differential equations (Hindmarsh et al., 2005), with "methods for stiff and nonstiff systems", was used for model integration. The time histories of the input and output signals are shown in Figure 2.7.



a) b)

Figure 2.7 - Time histories of a) the input force (N) and b) the mass position (m) for the system in Figure 2.3. The blue points are simulated data points, and the red line is the trajectory produced by the model with the nominal values for the four parameters.

After creating the simulated data, it was necessary to reoptimize the four free parameters to the data to ensure that the optimum set of parameter values was reached; even though the

parameter values used to create the data were known, a slightly different set of values would generally fit the data with a lower objective function value. The parameter identification functionality built into PottersWheel was used, and the "trust region" optimization algorithm was selected for this optimization, starting at the nominal parameter values used to create the data. Optimization was conducted in logarithmic parameter space. From the identified parameter values, the profile likelihood approach was run. As before, the CVODES solver was used for integration, and the trust region optimization algorithm was used to fit parameters in logarithmic space; the parameter bounds in Table 2.1 were applied during these optimizations as well. The χ^2 threshold for identifiability was calculated based on a simultaneous confidence level of 68% for which all four parameter confidence intervals hold jointly. For a normal distribution, a "68%" confidence interval covers plus-or-minus one standard deviation. Simultaneous confidence intervals consider the joint effects of parameter uncertainty on model validity. The resulting profile likelihood plots for each parameter are shown in Figure 2.8; the confidence interval values, true values, and estimated values for these parameters are listed in Table 2.2.

The profile likelihood results affirm the analytical results obtained in Section 2.3.1.1. Likelihood-based confidence intervals with finite upper and lower bounds indicate that the mass and damping coefficient parameters are identifiable from the information available. The points at which the profile likelihood crosses the upper threshold determine the upper and lower 68% simultaneous confidence interval values. (The differences seen in the quadratic approximation (gray) are attributed to the estimation of the curvature.) The flat profile likelihood plots with no local minimum indicate that the spring constants are both structurally unidentifiable. In other words, as each spring rate is incremented in increasing

and decreasing direction around its estimated value, it is possible to reoptimize the remaining parameter values such that no change in the objective function value is achieved. Namely, the two spring constant parameters are able to be adjusted such that a change in one value can be offset in a change by the other, i.e., to maintain $k_1 + k_2 = 4.4$. The functional relation between the two spring constants can be observed over the range of parameter values examined during the analysis. Plots created by PottersWheel show this relation in Figure 2.9.

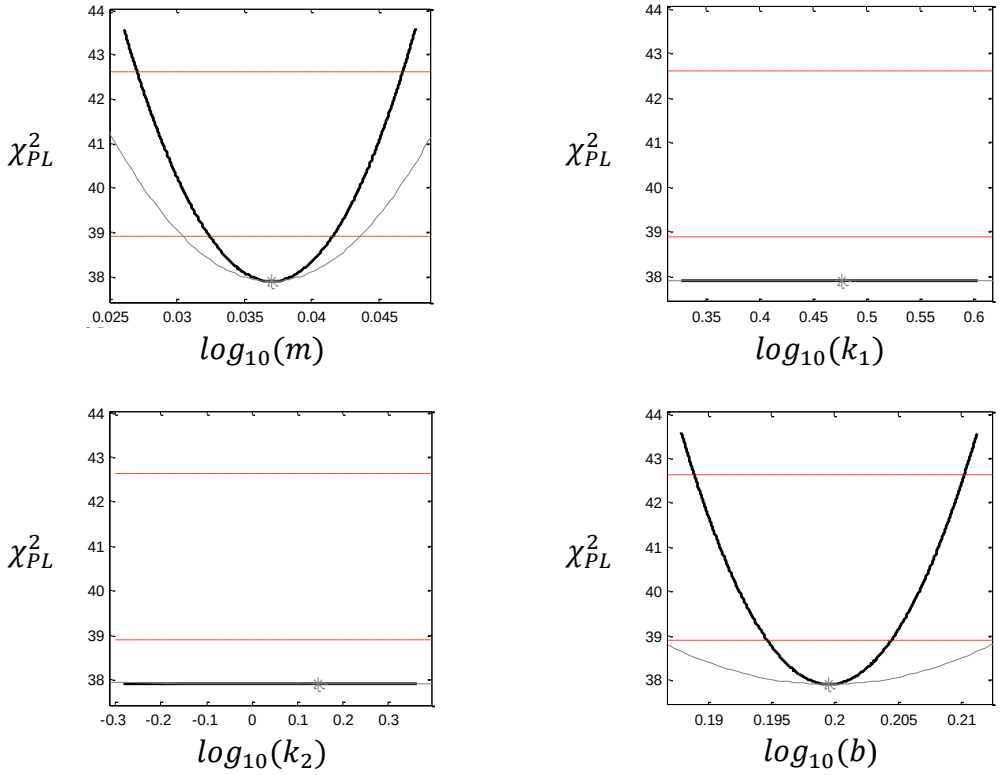


Figure 2.8 - Profile likelihood plots for the four parameters of the mass-spring-damper system in Figure 2.3, plotted in logarithmic space, for the simulated data shown in Figure 2.7. Black lines represent the profile likelihood; gray parabolas represent the quadratic approximation for asymptotic intervals. Gray asterisks at the valley of each curve indicate the estimated values of the parameters. The upper red dashed line of each plot represents the threshold for 68% simultaneous confidence intervals. The lower red dashed line represents the threshold for 68% pointwise confidence intervals. The mass and damping coefficient plots reach the upper threshold and are identifiable; the spring constant plots are flat and have no local minimum, so they are structurally unidentifiable.

Table 2.2 - True values of each parameter, as well as estimated values and 68% simultaneous likelihood-based confidence intervals (all in normal parameter space) for the simulated data shown in Figure 2.7.

Parameter	Units	θ_i^*	$\widehat{\theta_i}$	$CI_i^{-,PL}$	$CI_i^{+,PL}$
m	kg	1.1	1.09	1.06	1.11
k_1	N/m	3.0	3.00	0	$+\infty$
k_2	N/m	1.4	1.40	0	$+\infty$
b	N-s/m	1.6	1.58	1.54	1.62

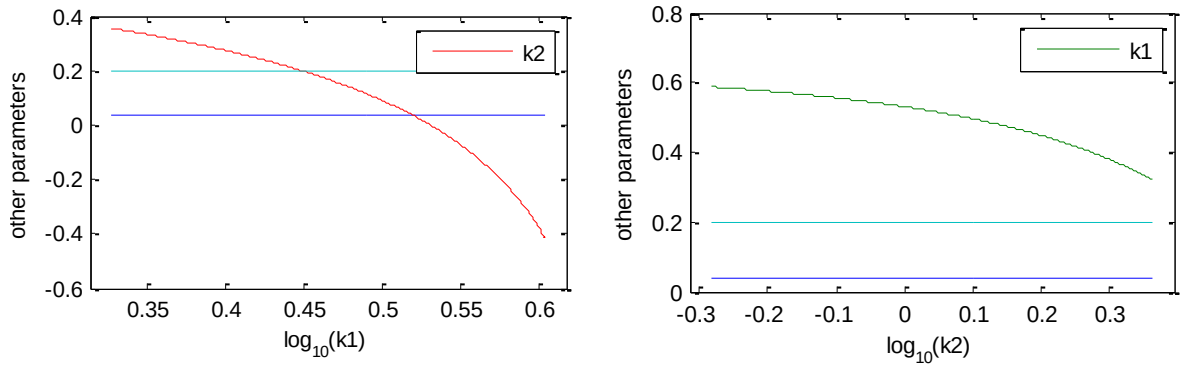


Figure 2.9 - Parameter values (all in logarithmic space) during profile likelihood analysis. The functional relation between structurally unidentifiable parameters k_1 and k_2 is visualized by changes in their values that maintain $k_1 + k_2 = 4.4$. The light blue and dark blue lines represent the damping coefficient and mass, respectively, which are unaffected by the relation.

As a second example of the profile likelihood identifiability approach, consider again the structurally identifiable mass-spring-damper system in Figure 2.2, modeled by the second-order equation,

$$m\ddot{x} + b\dot{x} + kx = f(t) \tag{2.26}$$

This model was implemented into PottersWheel in a manner similar to the previous example, with nominal values of $m = 1.1$, $k = 4.4$, and $b = 1.6$ and the same lower and upper bounds of 0.01 and 5.00 for each parameter. The same step input force of 40 N was used as well.

Although the 5 Hz sampling frequency on the mass position was retained, the Gaussian error model was modified to produce noise with a larger spread, i.e., a larger standard deviation. The time histories of the input and output signals are shown in Figure 2.10.

As in the previous example, the model parameters were first optimized to the dataset, and the profile likelihood analysis was run. The results of this analysis affirmed the structural identifiability of the model, and the detection of finite upper and lower confidence intervals for each of the three model parameters indicated their practical identifiability for the dataset given. As shown in Table 2.3, the presence of noise in the dataset widened the confidence intervals for the mass and damping coefficient compared to the values obtained in the previous example.

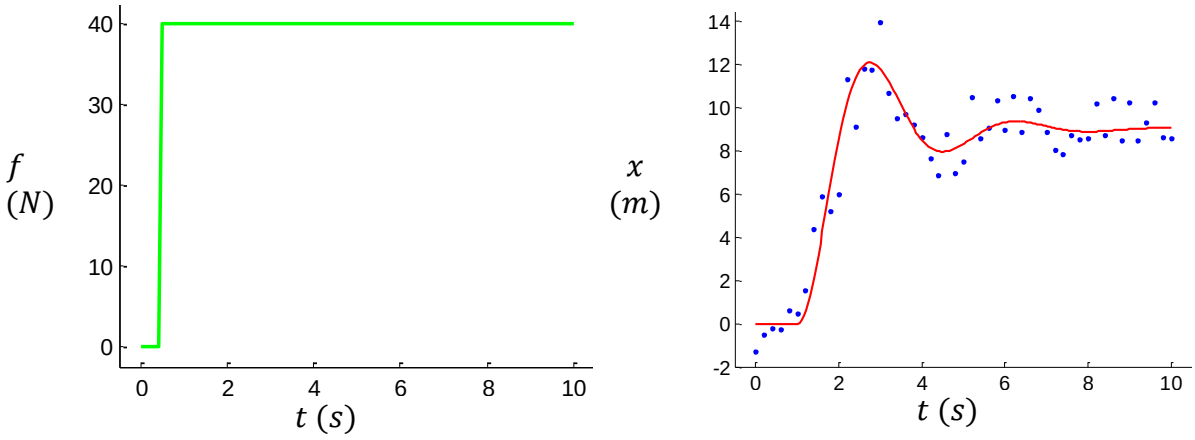


Figure 2.10 - Time histories of a) the input force (N) and b) the mass position (m) for the system in Figure 2.2. The blue points are simulated data points, and the red line is the trajectory produced by the model with the nominal values for the three parameters.

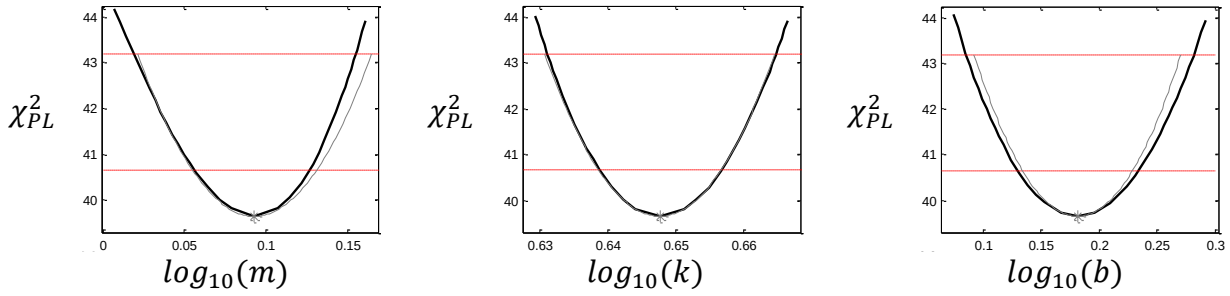


Figure 2.11 - Profile likelihood plots for the three parameters of the mass-spring-damper system in Figure 2.2, plotted in logarithmic space, for the simulated data shown in Figure 2.10. Black lines represent the profile likelihood; gray parabolas represent the quadratic approximation for asymptotic intervals. Gray asterisks at the valley of each curve indicate the estimated values of the parameters. The upper red dashed line of each plot represents the threshold for 68% simultaneous confidence intervals. The lower red dashed line represents the threshold for 68% pointwise confidence intervals. The model parameters' profile likelihoods each reach the upper threshold and are identifiable.

Table 2.3 - True values of each parameter, as well as estimated values and 68% simultaneous likelihood-based confidence intervals (all in normal parameter space) for the simulated data shown in Figure 2.10.

Parameter	Units	θ_i^*	$\widehat{\theta}_i$	$CI_i^{-,PL}$	$CI_i^{+,PL}$
m	kg	1.1	1.24	1.04	1.43
k	N/m	4.4	4.44	4.28	4.62
b	N-s/m	1.6	1.52	1.22	1.91

However, a confidence range has more meaning when translated into a range of physical performance. Figure 2.12 shows how the parameter confidence intervals translate into a family of trajectories for the mass position, forming a region of confidence in the output.

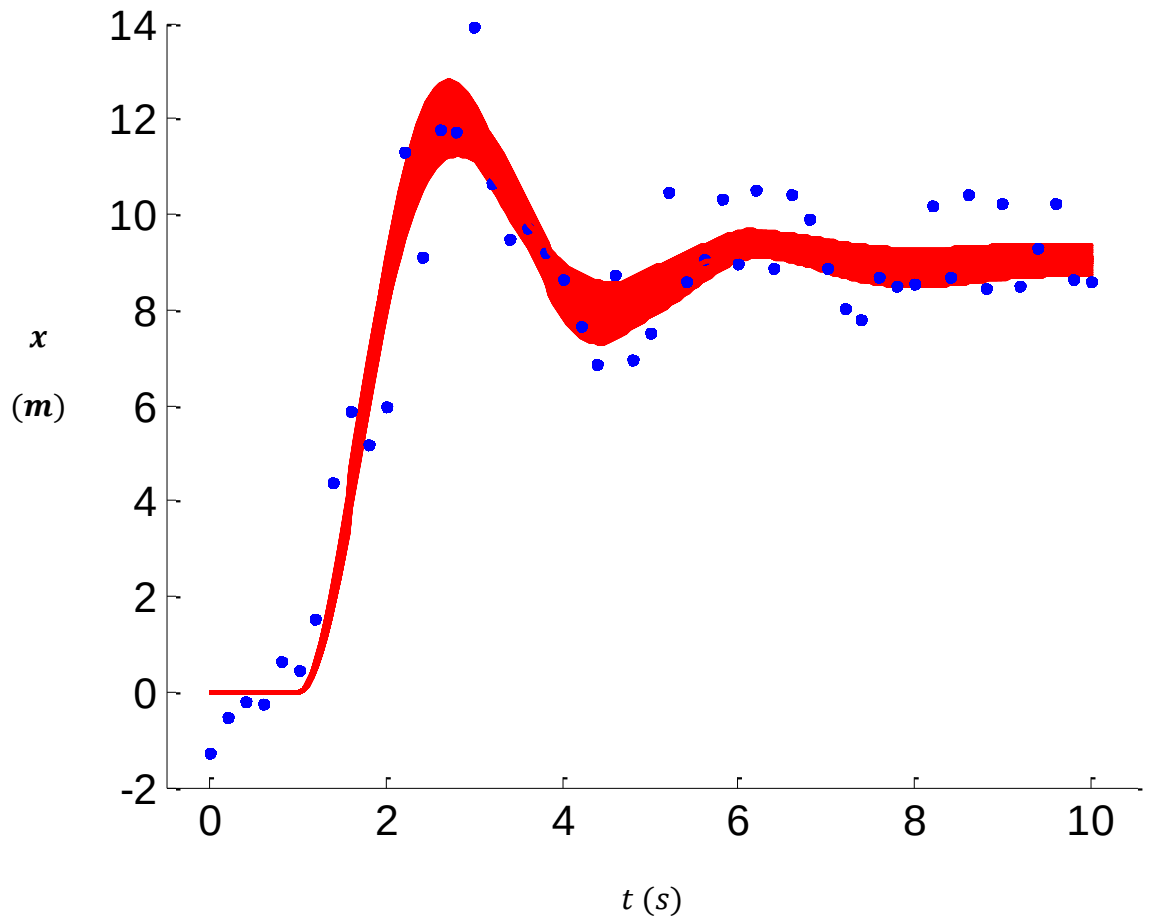


Figure 2.12 - Variation of the simulated trajectory (red) of the mass position within the parameter confidence intervals for the dataset given (blue).

References

Alasty, A., and A. Ramezani. (2002). Genetic algorithm based parameter identification of a nonlinear full vehicle ride model. *SAE Paper No. 2002-01-1583*.

Arikan, K. B. (2008). Identification of handling models for road vehicles. *PhD thesis*. Middle East Technical University.

Balsa-Canto, E., and J. R. Banga. (2010). Computational procedures for model identification. In: Systems Biology for Signaling Networks. S. Choi (ed.), 111-137. New York: *Springer.*

Banks, J. (1998). Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice. *John Wiley & Sons, Inc.*

Bellman, R., and K. J. Åström. On structural identifiability. *Mathematical Biosciences*, 7: 329-339.

Bernard, J., J. Greuning, and K. Hoffmeister. (1998). Evaluation of vehicle/driver performance using genetic algorithms. *SAE Paper No. 980227*.

Bernard, J., and O. Balling. (2004). Development of rollover maneuvers using optimization techniques. *SAE Paper No. 2004-01-2095*.

Bohlin, T. (2006). Practical Grey-box Process Identification. *Springer.*

Bolhasani, M. R., and Sh. Azadi. (2002). Parameter estimation of vehicle handling model using genetic algorithm. *SAE Paper No. 2002-01-1577*.

Book, R. (1996). Hydraulic system parameter identification using genetic algorithm methods. *In: Proceeedings of the Artificial Neural Networks in Engineering Conference*, 6: 351-356.

Book, R. S., and C. E. Goering. (2000). A new traction model for crawler tractors. *Transactions of the ASABE*, 43(1): 39-46.

Fittanto, D. A., and J. Puig-Suari. (2000). Utilizing a genetic algorithm to optimize vehicle simulation trajectories: determining initial velocity of a vehicle in yaw. *SAE Paper No. 2000-01-1616*.

Fujita, K., N. Hirokawa, S. Akagi, S. Kitamura, and H. Yokohata. (1998). Multi-objective optimal design of automotive engine using genetic algorithm. *In: Proceedings of ASME Design Engineering Technical Conferences*. Sept. 13-16, Atlanta, GA.

Goldberg, D. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning. *Addison Wesley*.

Goldberg, D. (1994). Genetic and evolutionary algorithms come of age. *Communications of the ACM*, 37(3): 113-119.

Grewal, M. S., and K. Glover. (1976). Identifiability of linear and nonlinear dynamical systems. *IEEE Transactions on Automatic Control*, 21(6): 833-837.

Grujicic, M., H. Marvi, G. Arakere, and I. Haque. (2010). A finite element analysis of pneumatic-tire/sand interactions during off-road vehicle travel. *Multidiscipline Modeling in Materials and Structures*, 6(2): 284-308.

Hengl, S., C. Kreutz, J. Timmer, and T. Maiwald. (2007). Data-based identifiability analysis of non-linear dynamical models. *Bioinformatics*. 23: 2612-2618.

Hindmarsh, A. C., P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward. (2005). SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software*, 31: 363-396.

Hoffmeister, K., and J. Bernard. (1998). Tread pitch arrangement optimization through the use of a genetic algorithm. *Tire Science & Technology*, 26(1): 2-22.

Jiafan, Z., Z. Yonglin, and Y. Qinghua. (2010). Numerical approach to parameter identifiability of mechanical system nonlinear parametric models. *2010 2nd International Conference on Industrial Mechatronics and Automation*, 16-19.

Karkee, M., and B. L. Steward. (2010). Local and global sensitivity analysis of a tractor and single axle grain cart dynamic model. *Biosystems Engineering*, 106(4): 352-366.

Ljung, L. (1999). System Identification: Theory for the User. Second Edition. Englewood Cliffs, NJ: *Prentice Hall*.

Maclay, D., and Dorey, R. (1993). Applying genetic search techniques to drivetrain modeling. *IEEE Control Systems*, 13(3): 50-55.

Maiwald, T., and J. Timmer. (2008). Dynamical modeling and multi-experiment fitting with PottersWheel. *Bioinformatics*, 24: 2037-2043.

Meeker, W. Q., and L. A. Escobar. (1998). Statistical Methods for Reliability Data. *John Wiley & Sons, Inc.*

Nocedal, J, and S. J. Wright. (1999). Numerical Optimization. New York: *Springer*.

Raue, A., C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15): 1923-1929.

Raue, A., C. Kreutz, T. Maiwald, U. Klingmüller, and J. Timmer. (2011). Addressing parameter identifiability by model-based experimentation. *IET Systems Biology*, 5(2): 120-130.

Saltelli, A., K. Chan, and E. M. Scott. (2000). Sensitivity Analysis. West Sussex, England: *John Wiley & Sons, Ltd.*

Serban, R., and J. S. Freeman. (2001). Identification and identifiability of unknown parameters in multibody dynamic systems. *Multibody System Dynamics*, 5: 335-350.

Vardeman, S. B., and J. M. Jobe. (2001). Basic Engineering Data Collection and Analysis. Pacific Grove, CA: *Duxbury*.

Venkataraman, P. (2009). Applied Optimization with MATLAB Programming. Second Edition. Hoboken, NJ: *John Wiley & Sons, Inc.*

Walter, E., and L. Pronzato. (1990). Qualitative and quantitative experiment design for phenomenological models–A survey. *Automatica*, 26(2): 195-213.

Walter, E., and L. Pronzato. (1997). Identification of Parametric Models from Experimental Data. *Springer*.

CHAPTER 3. IDENTIFIABILITY ANALYSIS OF A TRACTOR AND SINGLE AXLE TOWED IMPLEMENT MODEL

A paper to be submitted to *Biosystems Engineering*.

Simon L. Nielsen, Brian L. Steward

Abstract

The growing trend of model-based design in off-road vehicle engineering requires models that are sufficiently accurate for their intended application if they are to be used with confidence. Uncertain model parameters are often identified from measured data collected in experiments by using an optimization procedure, but it is important to understand the limitations of such a procedure and to have methods available for assessing the uniqueness and confidence of the results. A numerical approach based on the profile likelihood of parameters was utilized to evaluate the local structural and practical identifiability of a tractor and single axle towed implement model with six uncertain tire force model parameters from tractor yaw rate and implement yaw rate data. The analysis first considered datasets generated from simulation of the model with known parameter values to examine the effect of measurement error, sampling rate, and input signal type on the identifiability. The results showed that the accuracy and confidence of identification tended to decrease as quality and quantity of data decreased, to the point that several of the parameters were considered practically unidentifiable from the information available. The profile likelihood plots also indicated potential opportunities for model reduction. Second, the analysis considered the identifiability of the model from two datasets collected during field experiments, and the results again indicated parameters that were practically unidentifiable from the information

available. Overall, the study showed how different experimental factors can affect the amount of information available in a dataset for identification and that error in the measured data can propagate to error in parameter estimates.

Keywords: identifiability, parameter identification, optimization, experimental design, tractor and implement model, model-based design

3.1 Introduction

Off-road vehicle design and manufacturing companies are continually striving to meet customer needs by providing higher-quality and higher-performance products more quickly and at a lower cost. Advances in computer technology have had a major impact on engineering design and analysis over the last few decades (Dieter and Schmidt, 2009). An array of software packages for modeling and simulation of physical systems have been developed to take advantage of faster and more-flexible computing platforms (Åström et al., 1998). These technologies have fostered the growing trend of model-based design strategies in the off-road vehicle industry (Prabhu, 2007). In general, a model-based approach utilizes characterizations of system behavior to meet specified design requirements (Wymore, 1993). Model-based design has the potential to reduce reliance on physical prototypes, which can lead to time and cost savings during development (Prabhu, 2007; Lennon, 2008).

However, an ongoing limitation in the advancement of model-based design has been the development of accurate models in which one can put confidence regarding their ability to characterize a system (Radhakrishnan and McAdams, 2005). Without sufficient confidence, the usefulness of a model is restricted, and there will be hesitancy to rely on it to drive decision-making in design. Validation processes can be conducted to ensure that a

model can satisfactorily represent a physical system, at least within certain scenarios of interest (Ljung, 1999). Areas of concern include: the appropriateness of the model for the application, the accuracy of the mathematical representation of the model, and the accuracy of the model parameters (Bernard and Clover, 1994).

Off-road vehicle dynamics models are often mathematical models developed based on the principles of on-road vehicle dynamics, which can be found in e.g., Gillespie (1992) and Wong (2008). Off-road vehicle models have been developed for applications such as guidance controller design (Karkee and Steward, 2010a), traction modeling (Book and Goering, 2000), ride evaluation (Ahmed and Goupillon, 1997), handling evaluation (Previati et al., 2007), and real-time driving simulators (Fales, et al., 2005; Hummel et al., 2005; Karimi and Mann, 2006; Karkee et al., 2009). Depending on its level of fidelity, a vehicle model will typically incorporate a set of parameters to describe the physical system, including mass and inertia properties, geometric values, and other component and system properties such as stiffness or damping characteristics. Some of the parameter values may be uncertain due to the difficulty or impossibility of direct measurement, and certain parameter values that characterize a system well in one set of conditions may not be as appropriate as conditions vary (Kiencke and Nielsen, 2005; Karkee and Steward, 2011). Sensitivity analysis can be used to determine the effects of parameter variation on the model output (Jang and Han, 1997) as well as guide efforts to improve the certainty of specific parameters (Karkee and Steward, 2010b). In other words, variation in parameter values is used to evaluate variation in the output. Additionally, identification approaches can be used to determine vehicle model parameter values by finding the set of values for which the model output most closely represents the actual system output for a given input (Kiencke and

Nielsen, 2005). There are many examples of vehicle model parameter identification in the literature. Closeness of representation may be determined by comparing the time history of the same outputs and minimizing the error between them using an optimization approach. However, the values that minimize the error between a model structure and experimental data may differ from values obtained from other experimental approaches or other identified models.

For off-road vehicle models, the interaction of tires and soil is complex and difficult to characterize accurately (Wong, 1989). In particular, as noted by Karkee (2009), it is difficult to find a widely-accepted tire-soil model for lateral force development, which plays a primary role in steering response and yaw dynamics. However, researchers have used the well-known slip-angle-based tire model from on-road vehicle dynamics to relate tire slip angle to force development in off-road cases as well (Metz, 1993; Bevly et al., 2002; Karimi and Mann, 2006; Karkee and Steward, 2011). In some cases, values for the tire model parameters have been identified from vehicle-level data obtained during field experiments. Bevly et al. (2002) and Karimi and Mann (2006) each used tractor yaw rate data measured with a gyroscopic sensor along with front wheel steering angle data to identify cornering stiffness and relaxation length parameters of the front and rear tires of a linear bicycle model. Karkee and Steward (2011) used tractor yaw rate and heading angle data measured from Global Positioning System (GPS) receivers along with front wheel steering angle data to identify cornering stiffness and tire relaxation length parameters of a linear bicycle model of a tractor and single axle towed implement. In each of these cases, the difficulty in obtaining confident estimates for the tire model parameters was noted.

It is acknowledged that the off-road environment is less controlled than on-road environments, and the soil properties related to tire-soil interaction can vary substantially (Koolen and Kuipers, 1983; Crolla and El-Razaz, 1987; Karkee and Steward, 2011). Therefore, a model structure that assumes a constant value for the tire parameters is likely an insufficient representation of the physical system. However, without the ability to quantify this variation as a function of some other variable or system state, it is necessary to assume that, for each tire model parameter, a constant value exists that may minimize the error in the characterization. The term "cornering stiffness", referring to the slope of the lateral force versus lateral slip angle curve at zero slip, is actually a tire property that does not vary significantly for different surface conditions (Pacejka, 2006). The relationship between lateral force generation and slip angle in soil is nonlinear, so the use of this parameter in off-road studies to distinguish the force generation on different surfaces is a linear approximation that holds for a limited range (Metz, 1993). The results obtained in system-level parameter identification would not necessarily be expected to be the same as those obtained in controlled, lab-based tire tests, for example.

It is important to consider the possible limitations in parameter identification from experimental data. Although parameter sensitivity analysis lends insight into the effects that parameters have on the output, it does not necessarily show what effect the observed output (measured data) will have on the estimation of the parameters; that is, it does not show how variation in the output propagates to uncertainty in the estimated parameters. Model identifiability analysis is used to determine whether system measurements contain enough information to estimate the model parameters (Walter and Pronzato, 1997). Within this field, there are two subtypes frequently referred to as structural identifiability and practical

identifiability. Specifically, structural identifiability considers the mathematical structure of the model, independent of data, to determine if the parameters can be uniquely identified from the measured output (Walter and Pronzato, 1997). Several analytical methods are well-known for this analysis depending on the model type and size. With structural identifiability established, practical identifiability takes into account the properties of the measured data (Balsa-Canto and Banga, 2010), such as quantity, quality, and richness. Many practical identifiability studies make use of the Fisher Information Matrix, a measure of the precision of estimation based on the data at hand.

Despite its potential importance, a review of the literature shows that many mechanical system parameter identification studies do not seem to consider identifiability. Furthermore, if identifiability is considered, it will often only be structural in nature and will not consider the practical aspects of data collection. Unfortunately, identifiability analysis of complex linear models and nonlinear models using analytical methods is impractical, if not impossible, in many cases, even with the help of symbolic math computation (Arikan, 2008).

Identifiability analysis has been conducted in vehicle model identification studies (Serban and Freeman, 2001; Alasty and Ramezani, 2002; Arikan, 2008). Serban and Freeman (2001) noted the difficulty in applying global identifiability tests, but they developed a local, numerical test that determined if estimated parameters were at an "isolated minimum" of the optimization cost function. That test was demonstrated in the context of parameter identification of a multibody vehicle suspension model. Alasty and Ramezani (2002) tested the structural identifiability of a nonlinear, full-vehicle, ride model before using genetic algorithm optimization to identify 17 parameters from simulated data obtained from a high-fidelity multibody model. The model was linearized about an operating point to

determine the rank of the Jacobian of the Markov parameters, and identifiability of the linearized system was used to infer identifiability of the nonlinear system. Arikan (2008) examined the identifiability of a two degree-of-freedom linear vehicle handling model and a three degree-of-freedom nonlinear vehicle handling model prior to identification from data. The structural identifiability of the linear model was analyzed *a priori* using the transfer function approach for different observed output combinations and guided the sensor configuration for experimental data collection. The structural identifiability of the nonlinear model was examined using a differential algebra technique. The practical identifiability of the nonlinear model was examined based on the Fisher Information Matrix, which was used to ensure that there was not high correlation between parameters to be estimated. As noted by Arikan (2008), high correlation between parameters enables a change in one parameter value to be compensated by a change in another parameter value and limits identifiability.

Identifiability of dynamic models is an active topic of research in the field of systems biology. According to Raue et al. (2009), their reaction networks permit only a limited number of outputs to be measured, and experimental data is often of insufficient quantity and quality for parameter identification; furthermore, the size and complexity of their mathematical models often renders analytical identifiability methods inappropriate. The trend has been to utilize growing computational power to perform numerical identifiability analyses rather than use analytical approaches (Hengl et al., 2007; Raue et al., 2009). Raue et al. (2009) proposed a numerical approach for local identifiability analysis of arbitrary models by "exploiting" the profile likelihood of model parameters. The approach was able to detect structurally unidentifiable parameters due to functional relations and, since it was data-based, was able to detect practically unidentifiable parameters due to inadequate quality or

quantity of data. The approach considered identifiable parameters to have likelihood-based intervals in which the "true" value was to exist with a certain level of confidence.

In light of the challenges observed in identification of models, especially off-road vehicle tire force models, from data, it was desired to gain further insight into the feasibility of these approaches. In addition, it was desired to identify methods that could be used to determine the confidence/adequacy of parameter estimates, aside from system-level validation activities. The overall goal was to investigate the structural and practical identifiability of a tractor and single axle towed implement model, acknowledging that there are few methods widely demonstrated for this task. The purpose of this analysis was to better understand the influence of the model structure and the experimental conditions on the ability to identify certain parameters from data. The specific objectives were to:

- Examine the effect of measurement noise, data collection rates, and input excitations on the identifiability of tire model parameters from simulated data.
- Examine the identifiability of tire model parameters based on actual data collected in field experiments.

Therefore, of the three areas of concern in vehicle modeling discussed by Bernard and Clover (1994), this investigation is focused on the appropriateness of a model from a parameter identification standpoint. Results from this analysis may guide efforts to choose levels of model fidelity with parameters that can be reasonably identified from the experimental data available. The results may also provide information to guide experimental design for data collection.

3.2 Methods

In this study, identifiability analysis of a tractor and single axle towed implement model was performed based on the profile likelihood of the model parameters, using a numerical approach proposed by Raue et al. (2009). A dynamic bicycle model of the tractor-implement system (Karkee and Steward, 2010a) was used for the analysis. Several factors related to experimental data collection, including measurement noise, sampling rate, and system input characteristics were investigated to determine their influence on the identifiability of the model parameters of interest. The analysis was conducted first on simulated data with a specified noise model and then on data collected from field experiments.

3.2.1 Vehicle Model

The subject of this work was a tractor and single axle towed implement model of an agricultural tractor and grain cart system, studied extensively by Karkee and Steward (2010a). The actual system being modeled was a John Deere 7930 MFWD (mechanical front wheel drive) tractor (Deere and Co., Moline, IL) and a single axle, 18 m^3 (500 bu.) grain cart (model 500, Alliance Product Group, Kalida, OH). Those research efforts included the modeling of vehicle and tire force dynamics and an examination of open and closed loop system characteristics. Among the different models studied, they found that a dynamic bicycle model with tire relaxation length dynamics represented the system most accurately. This conclusion was based on a comparison of frequency response, and that model was used for the sensitivity analysis (Karkee and Steward, 2010b) and parameter identification studies (Karkee and Steward, 2011) that followed. Tire lateral forces were represented by a linear model based on the tire lateral slip angle, α, and a tire cornering stiffness, C_α, by,

$$F_y = -C_\alpha \alpha \tag{3.1}$$

The development of each tire slip angle was modeled as a first-order delay and parameterized by a relaxation length, σ, so that,

$$\dot{\alpha} = \frac{u}{\sigma}(\alpha_0 - \alpha) \tag{3.2}$$

where α_0 is the steady state slip angle (Bevly et al., 2002). The overall vehicle model is described by Eqs. (3.3) - (3.10):

$$\left(m^t + m^i\right)\dot{v}_c^t - m^i c\dot{\gamma}^t - m^i d\dot{\gamma}^i = -\left(m^t + m^i\right)u_c^t\gamma^t - C_{\alpha,f}^t\alpha_f^t - C_{\alpha,r}^t\alpha_r^t - C_{\alpha,r}^i\alpha_r^i \tag{3.3}$$

$$\left(I_z^t + m^i c^2\right)\dot{\gamma}^t - m^i c\dot{v}_c^t + m^i cd\dot{\gamma}^i = m^i cu_c^t\gamma^t - aC_{\alpha,f}^t\alpha_f^t + bC_{\alpha,r}^t\alpha_r^t + cC_{\alpha,r}^i\alpha_r^i \tag{3.4}$$

$$\left(I_z^i + m^i d^2\right)\dot{\gamma}^i - m^i d\dot{v}_c^t + m^i cd\dot{\gamma}^t = m^i du_c^t\gamma^t + (d+e)C_{\alpha,r}^i\alpha_r^i \tag{3.5}$$

$$\dot{\alpha}_f^t = \frac{v_c^t}{\sigma_f^t} + \frac{a\gamma^t}{\sigma_f^t} - \frac{u_c^t}{\sigma_f^t}\delta - \frac{u_c^t}{\sigma_f^t}\alpha_f^t \tag{3.6}$$

$$\dot{\alpha}_r^t = \frac{v_c^t}{\sigma_r^t} - \frac{b\gamma^t}{\sigma_r^t} - \frac{u_c^t}{\sigma_r^t}\alpha_r^t \tag{3.7}$$

$$\dot{\alpha}_r^i = \frac{v_c^t}{\sigma_r^i} - \frac{c\gamma^t}{\sigma_r^i} - \frac{(d+e)\gamma^i}{\sigma_r^i} + \frac{u_c^t}{\sigma_r^i}\varphi^t - \frac{u_c^t}{\sigma_r^i}\varphi^i - \frac{u_c^t}{\sigma_r^i}\alpha_r^i \tag{3.8}$$

$$\dot{\varphi}^t = \gamma^t \tag{3.9}$$

$$\dot{\varphi}^i = \gamma^i \tag{3.10}$$

and a schematic is shown in Figure 3.1. Full development of this model was documented by Karkee (2009). These equations can be represented in matrix differential equation representation as,

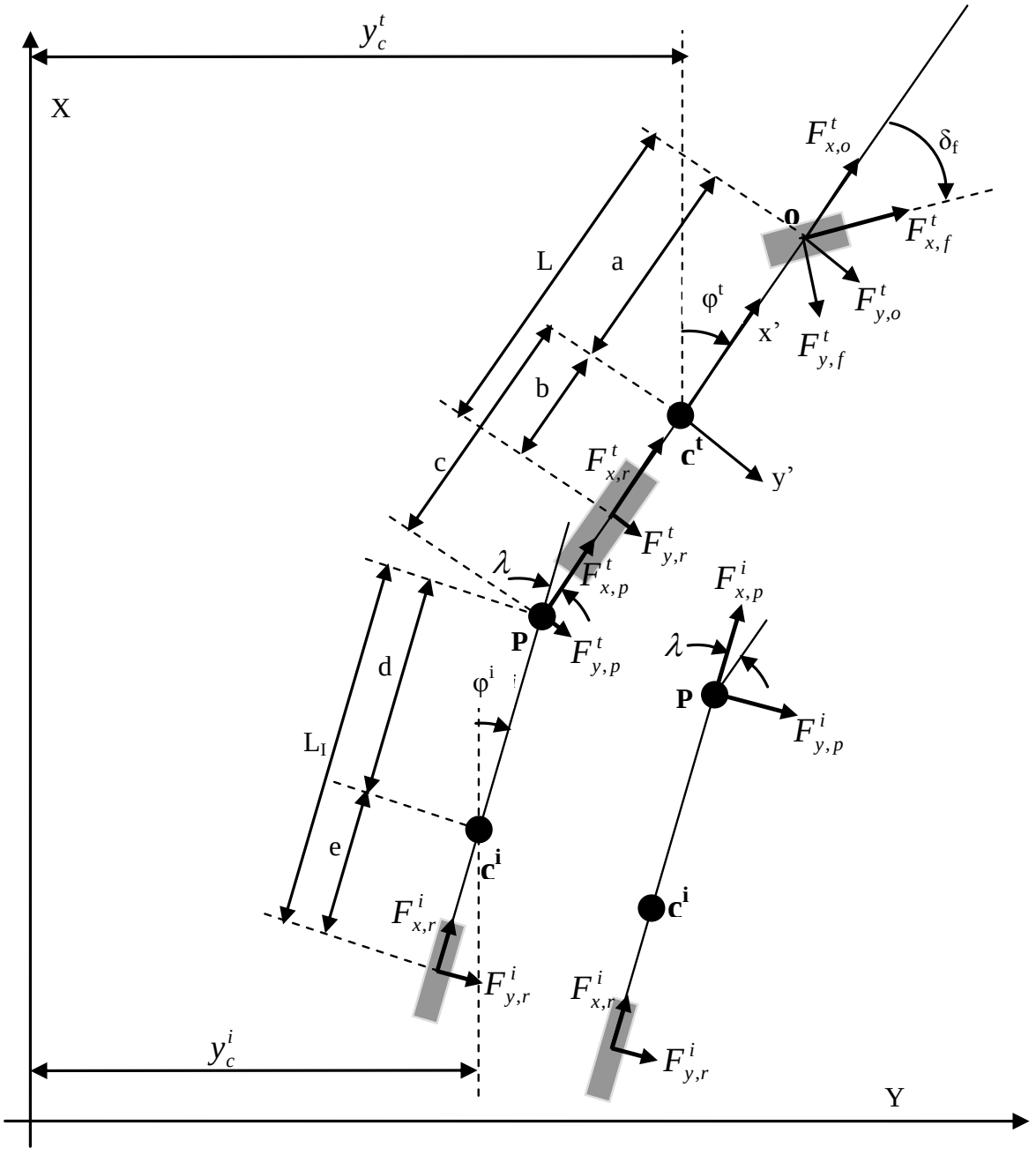
$$M\dot{X} = NX + PU \tag{3.11}$$

where the state vector is $X = \begin{bmatrix} v_c^t & \gamma^t & \gamma^i & \alpha_f^t & \alpha_r^t & \alpha_r^i & \varphi^t & \varphi^i \end{bmatrix}^T$, and the input is $U = [\delta]$. Matrices M, N, and P are given by,

$$M = \begin{bmatrix} m^t + m^i & -m^i c & -m^i d & 0 & 0 & 0 & 0 & 0 \\ -m^i c & I_z^t + m^i c^2 & m^i cd & 0 & 0 & 0 & 0 & 0 \\ -m^i d & m^i cd & I_z^i + m^i d^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.12}$$

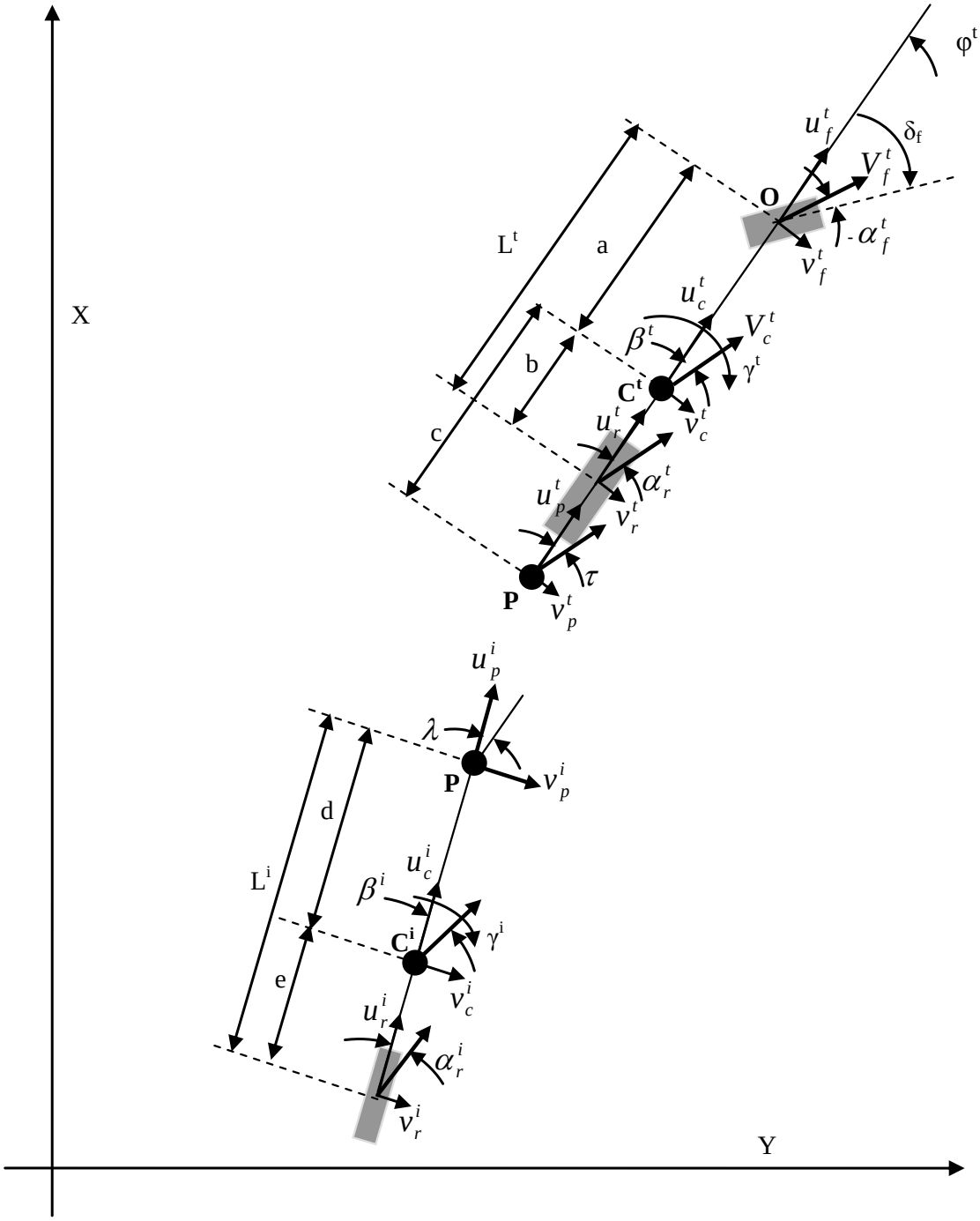
$$N = \begin{bmatrix} 0 & -(m^t + m^i)u_c^t & 0 & -C_{\alpha,f}^t & -C_{\alpha,r}^t & -C_{\alpha,r}^i & 0 & 0 \\ 0 & m^i c u_c^t & 0 & -aC_{\alpha,f}^t & bC_{\alpha,r}^t & cC_{\alpha,r}^i & 0 & 0 \\ 0 & m^i d u_c^t & 0 & 0 & 0 & (d+e)C_{\alpha,r}^i & 0 & 0 \\ \frac{1}{\sigma_f^t} & \frac{a}{\sigma_f^t} & 0 & -\frac{u_c^t}{\sigma_f^t} & 0 & 0 & 0 & 0 \\ \frac{1}{\sigma_r^t} & -\frac{b}{\sigma_r^t} & 0 & 0 & -\frac{u_c^t}{\sigma_r^t} & 0 & 0 & 0 \\ \frac{1}{\sigma_r^i} & -\frac{c}{\sigma_r^i} & -\frac{(d+e)}{\sigma_r^i} & 0 & 0 & -\frac{u_c^t}{\sigma_r^i} & \frac{u_c^t}{\sigma_r^i} & -\frac{u_c^t}{\sigma_r^i} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.13}$$

$$P = \begin{bmatrix} 0 & 0 & 0 & -\frac{u_c^t}{\sigma_f^t} & 0 & 0 & 0 & 0 \end{bmatrix}^T \tag{3.14}$$



a)

Figure 3.1 - Dynamic bicycle model of a tractor and single axled towed implement system (Karkee and Steward, 2010a); a) forces on the system, and b) velocities at different locations of the system (continued on text page).



b)

In the context of a field experiment for data collection, there are a limited number of system-level, tractor-implement outputs that can be reasonably measured with common, commercially-available sensors and data acquisition equipment and that have meaning with respect to the level of fidelity of the model being used. For the linear bicycle model considered here, these measurements could potentially include tractor and implement positions, heading angles, yaw rates, velocities, and accelerations. This study considered position and yaw rate measurements.

Vehicle positions are commonly measured using Global Positioning System (GPS) receivers, which may incorporate real-time kinematic (RTK) technology for increased accuracy. Although the GPS receiver will generally not be mounted directly over the tractor center of gravity (CG) in an experiment, it is assumed that this placement has been made possible for the purpose of this analysis. Based on the parameters and model states, the trajectory of the tractor CG was calculated as,

$$\dot{x}_c^t = u_c^t \cos(\varphi^t) - v_c^t \sin(\varphi^t) \tag{3.15}$$

$$\dot{y}_c^t = u_c^t \, sin(\varphi^t) + v_c^t \, cos(\varphi^t) \tag{3.16}$$

Similarly, the trajectory of the implement CG was calculated as follows based on the position of the tractor CG and the kinematics of the tractor and towed implement.

$$x_c^i = x_c^t - c \, cos(\varphi^t) - d \, cos\left(\varphi^i\right) \tag{3.17}$$

$$y_c^i = y_c^t - c \, sin(\varphi^t) - d \, sin\left(\varphi^i\right) \tag{3.18}$$

Yaw rates measurements are commonly obtained using gyroscopic sensors. These sensors can be mounted at any point on the object of interest as long as the measurement axis is oriented properly (i.e., parallel to an object's vertical axis). The tractor yaw rate, γ^t, and implement yaw rate, γ^i, were already calculated as states of the model.

3.2.2 Identifiability

As described by Walter and Pronzato (1997), physical systems are generally modeled in continuous time and described by a set of differential equations,

$$\dot{\boldsymbol{x}}(t) = f\big(\boldsymbol{x}(t), \boldsymbol{\theta}, \boldsymbol{u}(t)\big) \tag{3.19}$$

$$\boldsymbol{y}_m(t) = h\big(\boldsymbol{x}(t), \boldsymbol{\theta}, \boldsymbol{u}(t)\big) \tag{3.20}$$

where $\boldsymbol{x}$ is the state vector, $\boldsymbol{\theta}$ is the parameter vector, $\boldsymbol{u}$ is the vector of controlled inputs, t is time, and $\boldsymbol{y}_m$ is the vector of model outputs. For a common input vector, $\boldsymbol{u}$, the error vector, $\boldsymbol{e}_y$, between the system output vector, $\boldsymbol{y}$, and the corresponding model output vector, $\boldsymbol{y}_m$, is

$$\boldsymbol{e}_y(t, \boldsymbol{\theta}) = \boldsymbol{y}(t) - \boldsymbol{y}_m(t, \boldsymbol{\theta}) \tag{3.21}$$

These errors are also referred to as the "residuals". In an effort to obtain the best estimate of parameter values, $\widehat{\boldsymbol{\theta}}$, for the model to characterize the system, an objective function will be formulated that calculates a scalar value as a function of the output error, $\boldsymbol{e}_y$, and an optimization algorithm will be used to search the parameter space for the minimum value of the objective function. One of many optimization algorithms described in the literature can be used to search the parameter space for the optimal set of parameter values (Nocedal and Wright, 1999; Venkataraman, 2009).

As described by Walter and Pronzato (1997), identifiability refers to the uniqueness of a parameter vector $\widehat{\boldsymbol{\theta}}$ as an estimate of the true parameter vector $\boldsymbol{\theta}^*$ in a model M representing a physical system. Structural identifiability is considered independent of any data properties and considers a model to exactly represent the system of interest. Definitions for structural identifiability are given by e.g., Walter and Pronzato (1997) and Ljung (1999) and are closely described as follows. If the condition

$$M(\hat{\boldsymbol{\theta}}) = M(\boldsymbol{\theta}^*) \Rightarrow \hat{\theta}_i = \theta_i^* \tag{3.22}$$

holds for almost any $\boldsymbol{\theta}^*$ in $\mathbb{P}$ (the prior feasible set for $\boldsymbol{\theta}$; $\mathbb{P} = \mathbb{R}^n$ unless otherwise stated), then a parameter θ_i is classified as *structurally globally identifiable*. In other words, "identical input-output behavior" of two identical model structures implies that the estimated parameter set $\hat{\boldsymbol{\theta}}$ is unique and corresponds to the true parameter set $\boldsymbol{\theta}^*$. Furthermore, structural global identifiability of each parameter θ_i in $\boldsymbol{\theta}$ is a necessary condition for structural global identifiability of the model structure. The condition "almost any $\boldsymbol{\theta}^*$" functions to exclude atypical parameter values that may cause other parameters to become unidentifiable. If a model structure cannot be classified as globally identifiable, it may be possible to verify the model's local identifiability for some neighborhood $\mathbb{V}(\boldsymbol{\theta}^*)$ around the true parameter set. If Eq. (3.22) holds for $\hat{\boldsymbol{\theta}} \in \mathbb{V}(\boldsymbol{\theta}^*)$ then a parameter θ_i is classified as *structurally locally identifiable*. Each parameter θ_i in $\boldsymbol{\theta}$ must be at least structurally locally identifiable for the model structure to be classified as structurally locally identifiable. Consequently, local identifiability is a necessary condition for global identifiability. If there does not exist a neighborhood $\mathbb{V}(\boldsymbol{\theta}^*)$ for which Eq. (3.22) holds, then a parameter θ_i is classified as *structurally unidentifiable*. A model structure is structurally unidentifiable if one or more of its parameters is unidentifiable.

A number of methods are available for testing the structural identifiability of mathematical models (Walter and Pronzato, 1997). Methods for linear models are fairly well known but usually require the use of symbolic math analysis as model complexity increases. Methods for nonlinear models are generally more complicated. As noted by Serban and

Freeman (2001), "a global identifiability test is impractical for even the simplest models", so the scope must often be limited to local identifiability.

Practical identifiability, however, considers model identifiability in light of the characteristics of the experimental data used for parameter identification (Balsa-Canto and Banga, 2010), such as the quality, quantity, and richness of the data. Therefore, it is possible for a structurally identifiable parameter to be practically unidentifiable once experimental data is introduced. Quality refers to the presence of error in the output data. Quantity refers to the actual number of data points available; for data collection with respect to time, this will be determined by the sampling rate. Richness of data is related to the manner in which a system input is excited; richer datasets are generated by inputs that contain spectral content across the bandwidth of the model and persistently excite the system (Ljung, 1999). Statistical aspects of parameter identification have been described using maximum likelihood principles (Ljung, 1999).

Raue et al. (2009) and (2011) described a numerical approach to local structural and practical identifiability based on the profile likelihood of the model parameters. A detailed description of the approach can be found in those studies but is summarized as follows. For the optimization problem, they considered an objective function which is the weighted sum of squared residuals

$$\chi^2(\boldsymbol{\theta}) = \sum_{k=1}^{m}\sum_{l=1}^{d}\left(\frac{y_{kl}^D - y_k(\boldsymbol{\theta}, t_l)}{\sigma_{kl}^D}\right)^2 \tag{3.23}$$

where k is the index of m outputs measured, l is the index of d data points collected, y_{kl}^D is an experimental data point, y_k is a model output, and σ_{kl}^D is the corresponding measurement error of a data point. Assuming that the noise on the measurements is normally distributed,

$\epsilon \sim N(0, \sigma^2)$, minimization of this objective function yields maximum likelihood estimates of the parameter set, $\boldsymbol{\theta}$. Although asymptotic confidence intervals for the parameters can be obtained based on a quadratic approximation of the likelihood at the estimated parameter values if the model "sufficiently describes the experimental data", Raue et al. (2009) acknowledged that this approximation may not hold as well for cases with data of lower quality and/or quantity. For those cases, confidence intervals based on a "threshold" in the likelihood were recommended, defined by

$$\left\{\boldsymbol{\theta} | \chi^2(\boldsymbol{\theta}) - \chi^2\left(\hat{\boldsymbol{\theta}}\right) < \Delta_\alpha\right\} \tag{3.24}$$

$$\Delta_\alpha = Q\left(\chi^2_{df}, 1-\alpha\right) \tag{3.25}$$

where Δ_α is the $1-\alpha$ quantile of the χ^2-distribution with df degrees of freedom.

Raue et al. (2009) sought to efficiently search the parameter space around each parameter estimate by "exploring the parameter space for each parameter in the direction of least increase in χ^2". The profile likelihood was selected for that objective. This computation individually increments each parameter in increasing and decreasing directions around its estimate, reoptimizing all of the other parameters to the data and recording the χ^2 (objective function) value at each step. Therefore, the approach is able to capture the effects of parameter sensitivity as well as parameter interaction on the identification of model parameters. The computation produces a profile likelihood plot for each parameter, showing how its likelihood changes with respect to the parameter values. Based on Eqs. (3.24) and (3.25), upper and lower confidence bounds for a parameter are determined by the locations at which the likelihood crosses a certain χ^2 threshold.

The profile likelihood approach was used in this tractor-implement study instead of, for example, a quadratic likelihood approximation because it was not initially clear what constituted data of sufficient quantity and quality for this vehicle modeling application. Also, it was known that increased quantity and quality of data would cause the likelihood to converge toward the quadratic approximation anyway (Raue et al., 2011). Furthermore, this approach did not require any interaction with model equations and would thus be suitable for potential, future implementations with models of arbitrary format. The main drawback of the approach, being numerical in nature, was the computational requirements due to repeated function calls and optimization procedures during the analysis.

This profile likelihood approach is provided in the third-party PottersWheel mathematical modeling toolbox (Maiwald and Timmer, 2008) for MATLAB (The MathWorks, Inc., Natick, MA). Although the toolbox is tailored specifically toward the systems biology community, it has the capability to handle general mathematical models defined as a set of ordinary differential equations as well. In addition, the toolbox has many other functionalities that can be useful in mathematical modeling, parameter identification, and model analysis in many disciplines. Therefore, this toolbox was utilized to perform the identifiability analysis in this study.

3.2.3 Analysis

3.2.3.1 Simulated Data Analysis

The profile likelihood approach was first performed on simulated data for the tractor and single axle towed implement model described in Eqs. (3.3) - (3.10). Simulated data is sometimes analyzed in order to gain controlled insight into a model or procedure before considering experimental data, and it is generally advised to study properties of the

experiment before committing to experiments (Ljung, 1999). The intention was to analyze the model's identifiability free from any model characterization errors or unknown experimental error and to have complete control over the addition of error to the output data. These results would then represent a best-case scenario for parameter identification, upon which actual experimental data would not be likely to improve. Of primary interest in the simulated data analysis were the effects of measurement noise, data sampling rate, and input signal type on the identifiability of the tractor-implement model.

This effort was focused only on the identifiability of the tire model parameters, which were considered to be the most uncertain and most difficult to measure (Karkee and Steward, 2011). Although it is acknowledged that the values of the other parameters (masses, yaw moments of inertia, and geometric dimensions) have a degree of uncertainty associated with them as well, this assumption provided a narrowing of scope for the analysis. The values of these "fixed" parameters were measured or estimated by Karkee (2009) and are given in Table 3.1.

Table 3.1 - Dynamic bicycle model parameters for the JD 7930 tractor and Parker 500 grain cart system (Karkee, 2009).

Tractor			Implement (Grain Cart)		
Parameter	Nominal Value	Units	Parameter	Nominal Value	Units
a	1.7	m	d	3.62	m
b	1.2	m	e	0.1	m
c	2.1	m			
m^t	9391	kg	m^i	2127	kg
I_z^t	35709	kg-m^2	I_z^i	6402	kg-m^2

The PottersWheel toolbox required mathematical models to be entered in a specific format compatible with its functions. In particular, since the model was required to be

entered as a set of ordinary differential equations, it was necessary to convert the tractor-implement model from matrix differential equation representation, Eq. (3.11), to state-space representation to obtain the state equation,

$$\dot{X} = AX + BU \tag{3.26}$$

where $A = M^{-1}N$ and $B = M^{-1}P$. The MATLAB Symbolic Math Toolbox was used to perform the conversion. From this representation, the eight state equations were extracted. The only model input, the front wheel steer angle, δ, was specified using a driving input function with predefined input types. Fixed model parameters were specified directly, and the six unknown, or "free", tire model parameters were specified with a default value as well as minimum and maximum values for bounds. The nominal values of the tire model parameters were set at or near the values initially selected by Karkee and Steward (2010a) based on their review of the literature, but the upper and lower bounds, given in Table 3.2, were defined relatively wide around those nominal values, as if their values were unknown. These parameters are physically limited to real values greater than zero, so the lower bound selection was straightforward. However, the upper bounds were set more arbitrarily because there was no additional information available to guide their definition. The outputs for a given model were determined by the particular sensor configuration being simulated and were calculated based on the model states and parameters according to the development in Section 3.2.1. As part of the format for defining outputs, an error model with noise could be specified as well. Simulated data collection times were specified using a vector with start and stop times and intermediate times determined by a fixed collection frequency (e.g., 5 Hz).

Table 3.2 - Upper and lower bounds as well as nominal values for tire model parameters during the optimization.

Parameter	Units	Lower Bound	Nominal	Upper Bound
$C_{\alpha,tf}$	N/rad	10000	220000	700000
$C_{\alpha,tr}$	N/rad	10000	486000	700000
$C_{\alpha,ir}$	N/rad	10000	167000	700000
σ_{tf}	m	0.1	0.5	2.0
σ_{tr}	m	0.1	1.0	2.0
σ_{ir}	m	0.1	0.5	2.0

After loading a model into the PottersWheel graphical user interface (GUI), the option to create simulated data was used. The simulation used the nominal values of the free parameters and applied the specified error model to the outputs. The built-in CVODES solver for ordinary differential equations (Hindmarsh et al., 2005), with "methods for stiff and nonstiff systems", was used for integration.

After creating the simulated data, it was necessary to reoptimize the six free parameters to the data to ensure that the optimum set of parameter values was reached; even though the parameter values used to create the data were known, a slightly different set of values will generally fit the simulated data with a lower objective function value. PottersWheel was used for the parameter identification process. A "trust region" optimization algorithm was selected for this process, starting from the known parameter values used to create the data. A global optimization technique would generally be chosen for the initial optimization step of a complex, multi-dimensional identification problem, but it was assumed that the optimal parameter set for the simulated data could be reached with a local technique since it started at the known, true values. Optimization was conducted in logarithmic parameter space since the normal values of the parameters extended more than one order of magnitude and can only have positive values.

From the identified parameter values, the profile likelihood approach was run. As before, the CVODES solver was used for integration, and the trust region optimization algorithm was used to fit parameters in logarithmic space; the parameter bounds in Table 3.2 were applied during these optimizations as well. The χ^2 "threshold" for identifiability was calculated based on a simultaneous confidence level of 68% for which all parameter confidence intervals hold jointly. For a normal distribution, a "68%" confidence interval covers plus-or-minus one standard deviation. Simultaneous confidence intervals consider the joint effects of parameter uncertainty on model validity. The computation time required to complete the analysis depended on the model, the amount of data, and the particular configuration of the profile likelihood settings but was typically between five and ten minutes per parameter for conservative settings on a 2.8 GHz workstation with 8 GB of RAM. The profile likelihood was computed in relatively small steps to ensure that it would be smooth; this required a greater number of function calls. Sometimes, a slightly better optimum of the objective function was found during the computation, and it was necessary to rerun the analysis for that better set of parameter values.

3.2.3.1.1 Simulated Step Input Sampled at 5 Hz

The first case considered the influence of measurement noise on the identifiability of the six tire model parameters. Whether a function of sensor error, unmodeled dynamics, or any other unwanted corruption source, noise is a practical issue in the measurement of physical system outputs. The tractor forward velocity was held constant at 4.5 m/s, and a rate-limited step input from 0 to 10 degrees at the front wheels was applied over 0.5 seconds. The model outputs were the tractor yaw rate and the implement yaw rate, each sampled at 5 Hz for a period of 10 seconds, a length of time that provided measurements that were

composed of approximately half transient response and half steady-state response. The 5 Hz sampling frequency was selected based on the specifications of a GPS receiver with yaw rate sensing capabilities that is commonly used in agricultural applications. For the purpose of this investigation and to maintain the validity of the identifiability approach, Gaussian noise was added to the simulated data. The signal-to-noise ratio (SNR) of data has been expressed as the reciprocal of the coefficient of variation (Meeker and Escobar, 1998), in which

$$SNR = \frac{E(T)}{SD(T)} \tag{3.27}$$

In this equation, $E(T)$ and $SD(T)$ represent the expected value and standard deviation, respectively, of a continuous random variable T. In this study, the numerator term of the SNR equation was specifically defined as the maximum amplitude of each tractor yaw rate signal, $\max(\gamma^t)$, in the maneuver, such that

$$SNR = \frac{\max(\gamma^t)}{\sigma} \tag{3.28}$$

The denominator term, σ, was the standard deviation of the specific noise model applied to the output. Therefore, the signal-to-noise ratio was varied from 1000, a nearly undistorted signal, to 12.5, a signal for which the transient response was nearly impossible to detect visually. Within this range, SNR values of 100, 50, 25, and 16.67 were considered.

3.2.3.1.2 Simulated Step Input Sampled at 10 Hz

The second case considered the influence of sensor sampling rate on the identifiability of the six tire model parameters. Experiment 1 was repeated for the same 10 second period with a 10 Hz sampling rate and compared with the 5 Hz sampling rate results. The increased rate provided a better opportunity to capture the transient yaw rate response and doubled the amount of data available for a given time period.

3.2.3.1.3 Simulated Chirp Input Sampled at 5 Hz

The third case considered the influence of the input signal type on the identifiability of the six tire model parameters. Experiment 1 was repeated using a chirp steering input (a sine wave with a frequency varying linearly with time) instead of a rate-limited step input. The initial steering frequency was 0.1 Hz, the final steering frequency was 0.5 Hz, and the signal amplitude was 10 degrees. The frequency values were selected based on handwheel input rate limitations encountered in the actual system. The intention was to use an input signal that would persistently excite the system and produce output data composed entirely of transient response.

3.2.3.2 Experimental Data Analysis

The identifiability of the tractor-implement model from data collected during actual field experiments was investigated as well. The data used for the analysis were collected by Karkee (2009) in 2008 and 2009 as part of a parameter identification study using the tractor and implement system described in Section 3.2.1. In that study, tractor and implement CG (center of gravity) trajectories, heading angles, and yaw rates were measured at 5 Hz using agricultural GPS receivers with yaw rate sensors, and the front wheel steering angle was collected using a rotary potentiometer that was installed on the left-wheel kingpin by the tractor manufacturer. Data were collected in an agricultural field that had been planted to alfalfa three growing seasons before the experiments and had been uncultivated since that planting. Data were collected for a variety of steering maneuvers at three different forward velocities. To maintain the assumptions needed for the linear model, steering angles were limited to +/- 10 degrees, and the forward velocity was held approximately constant. The data were used to identify the six tire model parameters of the tractor-implement model.

Although the parameter estimation approach in that study was able to identify parameter values that improved the ability of the model to represent system behavior over initial parameter values, there was substantial uncertainty in some of the parameter estimates, and the suitability of the estimates varied between different experimental trials. Nonetheless, based on the results of the investigation by Karkee (2009), it is assumed that the eighth-order tractor-implement model presented in Section 3.2.1 provides a "sufficient" representation of the measured data and can be used in this identifiability study.

Two datasets from those field experiments were examined separately in this identifiability study using the same numerical approach applied to the simulated data in Section 3.2.3.1. The first dataset was collected while applying two step steering angle inputs to the tractor, and the second dataset was collected while applying a chirp steering angle input to the tractor. For each set, the data were uploaded into the PottersWheel toolbox, and the front wheel steering angle sensor data were used to drive the steering angle of the tractor-implement model. It was assumed that any noise in the steering angle sensor data was random noise of low amplitude about the true steering angle value and of a frequency high enough to have little to no impact on the output. The error values of the data points that were used to weight the residuals in the objective function calculation, Eq. (3.23), were estimated with respect to a cubic smoothing spline that was fit to the data.

Contrary to the situation in the simulated data analysis, the "true" values of the tire model parameters were unknown in the experimental cases, and it was more difficult to know if the best estimates of the parameters were obtained. Since the performance of local optimization techniques is dependent on the initial values of the unknown parameters, a global optimization algorithm was used first (outside of PottersWheel) to determine the best

estimates for the data available. A genetic algorithm (GA) optimization approach, which is a stochastic procedure that searches the parameter space in a heuristic manner (Goldberg, 1989), was used initially. The goal was to obtain a set of parameter values at or near the global optimum. The approach was implemented using the GA functions in the Global Optimization Toolbox in MATLAB which minimized an objective function that quantified the sum of squared error for the tractor yaw rate and implement yaw rate time histories. The values obtained using GA then served as a starting point for the trust region method in PottersWheel, which yielded another slight improvement in fit, subject to the upper and lower bounds specified in Table 3.2. The profile likelihood approach was then configured and run in a manner similar to that used in the simulated data analysis.

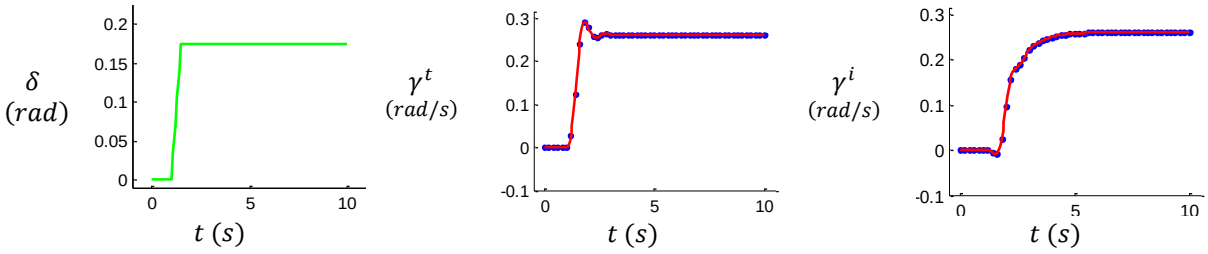
3.3 Results and Discussion

3.3.1 Simulated Data Analysis

Identifiability analysis of the simulated data with a known noise model provided a means by which the model structure could be evaluated. The impact of different conditions associated with experimental data collection became apparent as they were varied, and the overall trends agreed with expectations. Analysis of the results is based on the methods and identifiability definitions described by Raue et al. (2009) and (2011).

3.3.1.1 Simulated Step Input Sampled at 5 Hz

The nearly noise-free dataset with signal-to-noise ratio of 1000 represented the ideal situation for parameter identification, as shown in Figure 3.2. The profile likelihood of each parameter was nearly parabolic, as shown in Figure 3.3, approaching the quadratic approximation for asymptotic confidence intervals.



a) b) c)

Figure 3.2 - Time histories of a) front wheel steering input (rad) and simulated data with signal-to-noise ratio of 1000 for b) tractor yaw rate (rad/s), and c) implement yaw rate (rad/s).

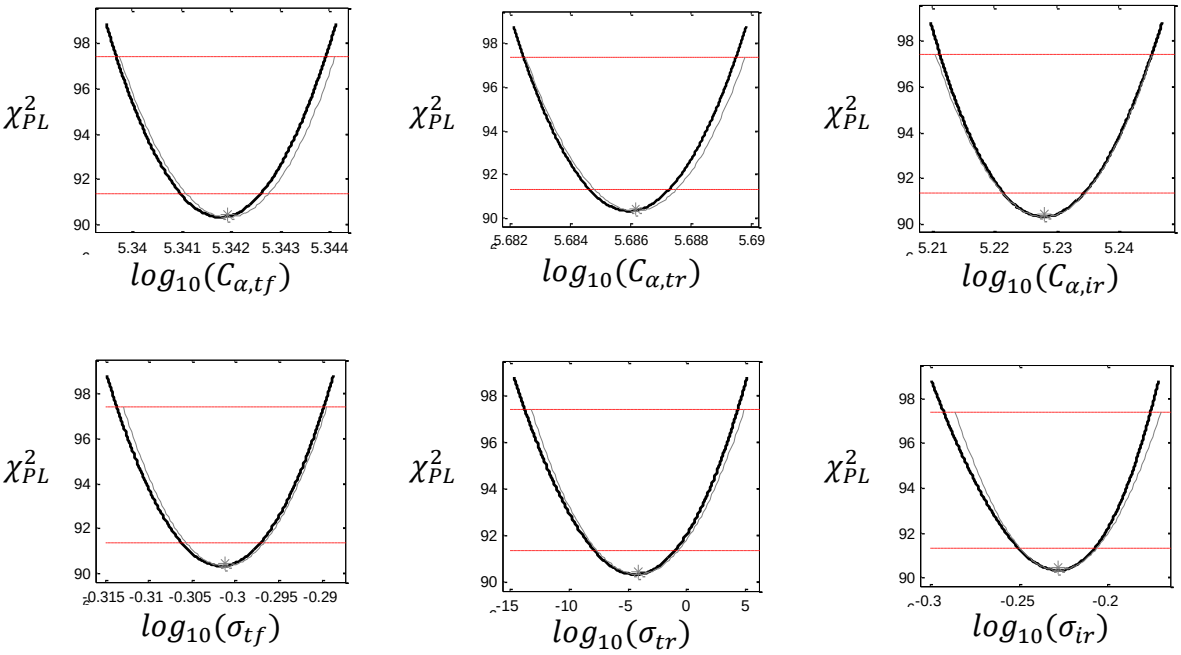


Figure 3.3 - Profile likelihoods for each of the six tire model parameters, plotted in logarithmic parameter space, for the simulated data shown in Figure 3.2. Black lines represent the profile likelihood; gray parabolas represent the quadratic approximation for asymptotic intervals. Gray asterisks at the valley of each curve indicate the estimated values of the parameters. The upper red dashed line of each plot represents the threshold for 68% simultaneous confidence intervals. The lower red dashed line represents the threshold for 68% pointwise confidence intervals.

Confidence intervals with finite upper and lower bounds indicated that the tractor-implement model, with six uncertain tire model parameters, was practically identifiable from the tractor

yaw rate and implement yaw rate data and, therefore, structurally identifiable. The values of these confidence intervals, as well as the true values and the estimated values, are listed in normal parameter space in Table 3.3.

Table 3.3 - True values of each parameter, as well as estimated values and 68% simultaneous likelihood-based confidence intervals (all in normal parameter space) for the simulated data shown in Figure 3.2.

Parameter	Units	θ_i^*	$\hat{\theta}_i$	$CI_i^{-,PL}$	$CI_i^{+,PL}$
$C_{\alpha,tf}$	N/rad	220000	219700	218600	220800
$C_{\alpha,tr}$	N/rad	486000	485400	481400	489200
$C_{\alpha,ir}$	N/rad	167000	169000	162700	175900
σ_{tf}	m	0.5	0.500	0.486	0.513
σ_{tr}	m	1.0	0.990	0.969	1.010
σ_{ir}	m	0.5	0.591	0.510	0.667

Even though it was deemed identifiable, the implement tire relaxation length was estimated least accurately, and its true value was narrowly outside of the likelihood-based confidence region. However, as the signal-to-noise ratio decreased from 1000, practical unidentifiabilities became apparent based on widening, sometimes infinite, confidence intervals and less accurate identification of the true parameter values. The first practically unidentifiable parameter to emerge from the analysis was the implement tire relaxation length for SNR = 100, which was fit to its lower bound of 0.1. Sensitivity analysis had already shown that the implement tire relaxation length was among the parameters to which the system dynamics were the least sensitive (Karkee and Steward, 2010b), so it was expected that it would be difficult to estimate this parameter confidently based on the system output. In fact, this parameter value was frequently fit to one of its bounds, which may be indicative of a poor fit due to inappropriate model structure. The inability of the implement tire relaxation length's profile likelihood to reach the threshold value for identifiability

showed that changes in that parameter caused little improvement in the model's fit to the data and/or could be masked by changes in other parameters such that there was little increase in the objective function value. From a model calibration standpoint, it is also more difficult to converge to the optimum value of a parameter if its effect on the output cannot be easily detected.

As the signal-to-noise ratio continued to decrease from 100 to 12.5, the next practical unidentifiabilities to emerge were the tractor's front and rear tire relaxation lengths, the implement tire cornering stiffness, and even the tractor's rear tire cornering stiffness. These results also follow the outcomes of the previously mentioned sensitivity analysis; that is, the parameters to which the system dynamics are most sensitive are also the ones that can be estimated most confidently from the output data. A compilation of the results is shown in Figure 3.4. The red bars represent likelihood-based, 68% simultaneous confidence intervals for each of the six tire model parameters. The number to the left of each bar indicates the signal-to-noise ratio of the data it pertains to. An arrow on the upper and/or lower end of a bar indicates a practical unidentifiability due to a confidence bound extending to +/- infinity (in logarithmic space). The black line in each cluster indicates the true value of the parameter which was used to create the simulated data. Each green diamond in a red bar indicates the estimated value of the parameter from the data.

The confidence intervals for each of the six parameters have more physical meaning once translated into a range of performance in the tractor-implement system. Within the confidence intervals for a particular signal-to-noise ratio, a family of trajectories for the tractor yaw rate and implement yaw rate could be determined as well as for the position of the tractor CG and the implement CG based on Eqs. (3.15) - (3.18).

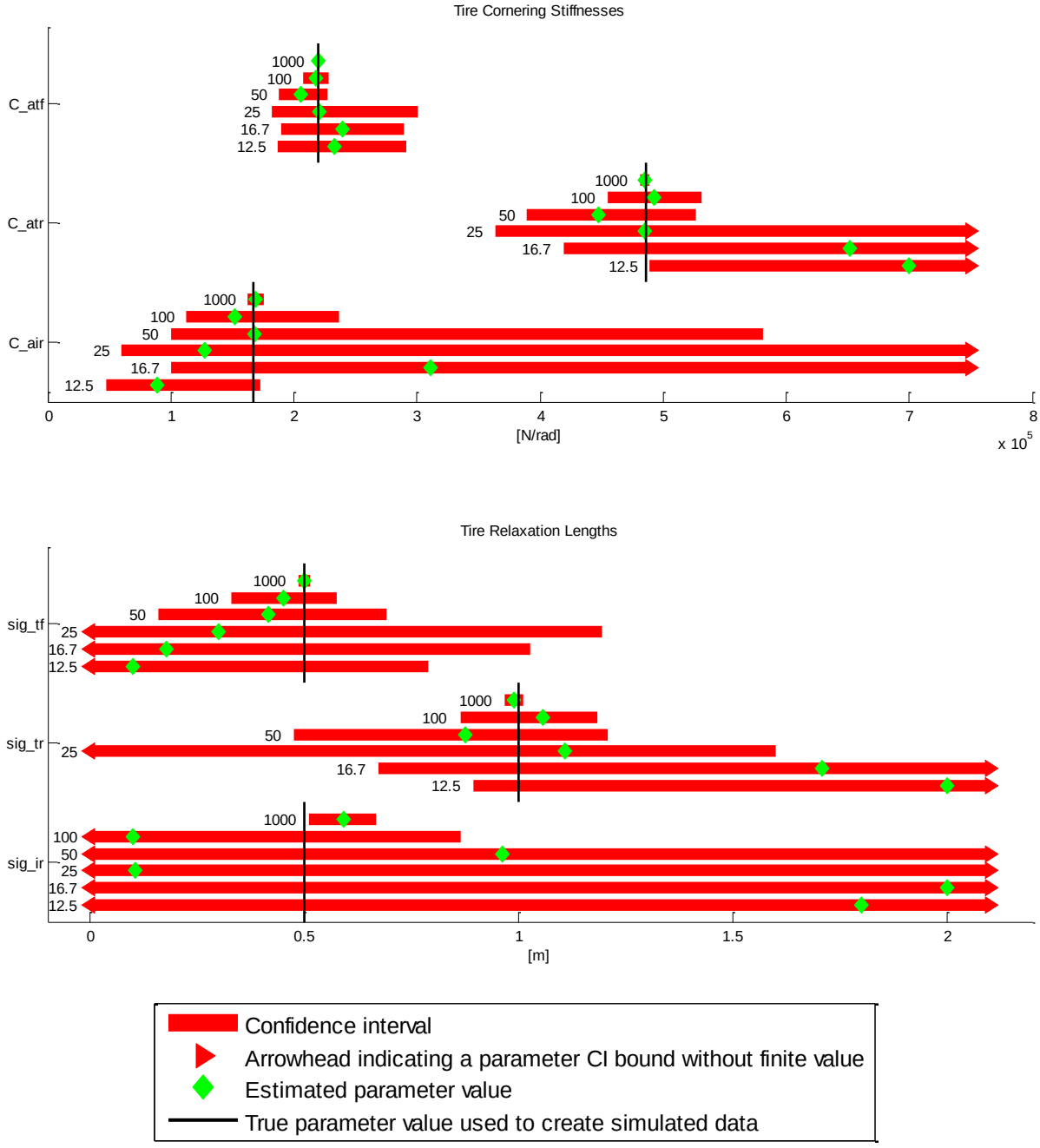


Figure 3.4 - Likelihood-based, 68% simultaneous confidence intervals (red bars) for each of the six tire model parameters in the first experiment of the simulated data analysis (5 Hz collection frequency of the tractor yaw rate and implement yaw rate for a rate-limited step steer input of 10 degrees). The number to the left of each bar indicates the signal-to-noise ratio of the data it pertains to.

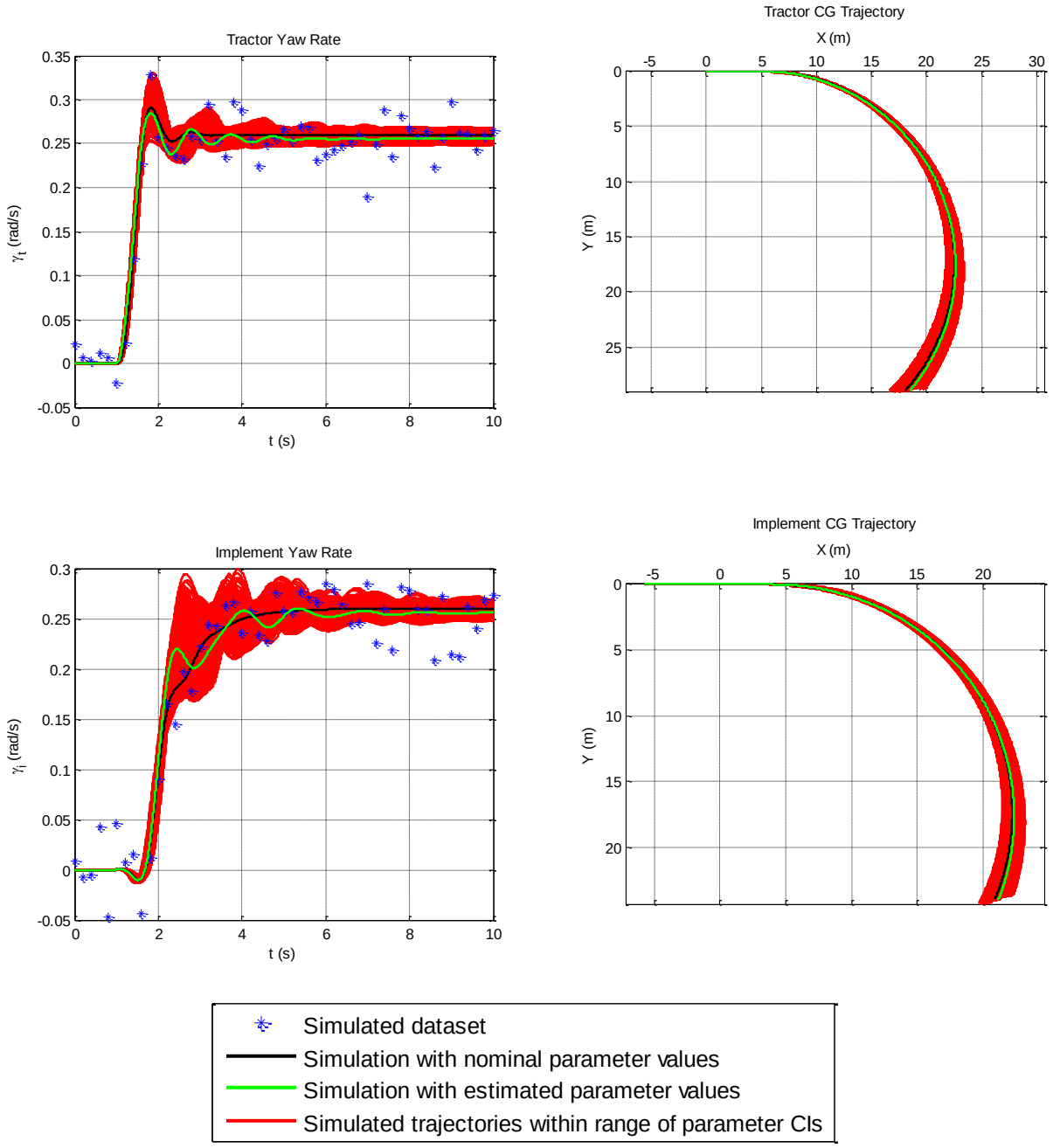


Figure 3.5 – Variation of the simulated trajectories within the 68% simultaneous confidence intervals obtained for each of the six tire model parameters, for the simulated rate-limited step input sampled at 5 Hz and with SNR = 12.5.

The plots in Figure 3.5 show these four signals for the noisiest case, where SNR = 12.5. The nominal and estimated CG trajectories matched very closely despite the large confidence

intervals for the parameters and the inability to closely estimate the true parameter values for all but the front tractor tire cornering stiffness.

As shown in Figure 3.6, the profile likelihood plots lent some additional insight into the analysis. As the signal-to-noise ratio decreased, the unidentifiability of the tire relaxation length parameters was generally manifested as a flattening or limited increase of the profile likelihood in the decreasing direction. This indicated the inability to determine a finite likelihood-based lower confidence bound based on the amount and quality of data available. As a tire relaxation length approaches its theoretical lower limit of $\sigma = 0$, the time constant associated with its slip angle dynamics decreases to zero as well, and the relaxation length dynamics are removed. Therefore, the inability to determine a lower confidence bound for these relaxation length parameters from this data suggests a growing inability to distinguish between the current dynamic model with relaxation length dynamics included and a reduced model with these dynamics removed. This result does not necessarily suggest that the reduced model is a better characterization of the actual system – it just suggests that it becomes more difficult to distinguish the adequacy of the reduced model versus the full model as the noise increases in the information available. It also calls into question the rationality of attempting to identify the relaxation length parameters from that information. ("Distinguishability" is, in fact, another property used to compare the suitability of two or more models (Walter and Pronzato, 1997).) The red dots in the plots indicate simulation points where at least one of the other parameter values was fit to one of its bounds. Discontinuities and smaller valleys encountered in the traversal of the profile likelihood are local minima (Raue et al., 2009).

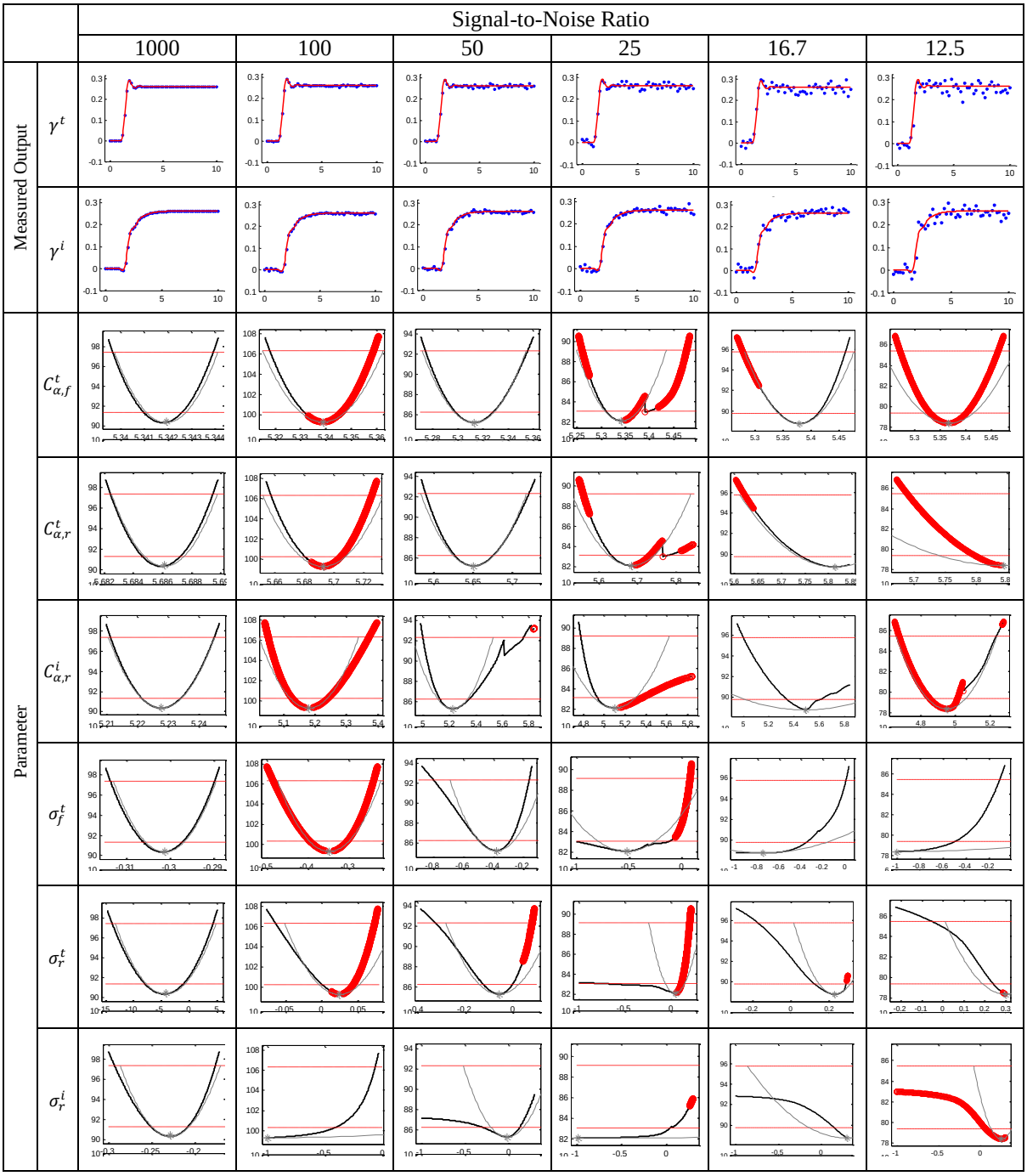


Figure 3.6 - Profile likelihood plots for each tire model parameter (in logarithmic space) when identified from simulated 5 Hz tractor yaw rate (rad/s) and implement yaw rate (rad/s) data for a 10 degree rate-limited step steer input. Simulations are grouped column-wise by the signal-to-noise ratio of the error applied to the measured outputs. Red dots on a profile likelihood plot indicate simulation points where at least one of the other five parameters was fit to one of its bounds. Discontinuities and smaller valleys encountered in the traversal of the profile likelihood are local minima.

Likewise, unidentifiability of the implement cornering stiffness and even the tractor's rear tire cornering stiffness for greater noise levels were generally manifested as a flattening out or limited increase of the likelihood in the increasing direction, which indicates the inability to determine a finite likelihood-based upper confidence bound (within the specified parameter range) based on the amount and quality of data available.

The outcome for the cornering stiffness parameters was more difficult to interpret, because the unidentifiabilities were determined with respect to the upper bounds of 700000 N/rad that were chosen somewhat arbitrarily; a higher upper bound or a lower confidence threshold could change this classification. Additional information about the practical range of these parameter values would help set these bounds with more certainty. However, as a tire cornering stiffness approaches its theoretical upper limit of $C_\alpha = \infty$, the tires will have no slip angle and will move in the direction that they are facing. Therefore, the inability to determine an upper confidence bound for the cornering stiffness parameters from this data and the flattening of the likelihood in the upper direction may indicate growing inability to distinguish between a dynamic model and a kinematic model. Again, this conclusion is only made with regard to the data used for identification.

If the effect of the towed implement on the tractor is neglected, the identifiabilities of the tractor's front and rear cornering stiffness values may be partially explained by a vehicle steering response characteristic known as the "understeer gradient", given by,

$$K = \frac{W_f}{C_{\alpha,f}} - \frac{W_r}{C_{\alpha,r}} \tag{3.29}$$

where W_f and W_r are the loads on the front and rear axle, respectively. The understeer gradient of a vehicle, measured in the units degrees/g, is an important property that indicates

how its steering angle must change as the lateral acceleration changes (Gillespie, 1992). Since the maneuvers in this analysis were conducted with constant forward velocities and relatively small steering angles, it can be assumed that longitudinal load transfer is negligible, and the tractor's understeer gradient should be approximately constant. Therefore, a change in one of the cornering stiffness values can be balanced by a change in the other value such that the understeer gradient steering response characteristic can be held constant.

It should also be noted that the steady-state portions of the implement yaw rate response do not provide information for the identification of the implement tire cornering stiffness values. In steady state conditions, the yaw rate of the implement is determined solely by the yaw rate of the tractor. This factor likely plays a role in the identifiability of the implement tire cornering stiffness.

3.3.1.2 Simulated Step Input Sampled at 10 Hz

Doubling the tractor yaw rate and implement yaw rate sampling rates from 5 Hz to 10 Hz resulted in slight improvements in both accuracy and confidence of identification compared to simulated case 1. Confidence intervals for identifiable parameters generated from the 10 Hz data tended to be narrower than their respective confidence intervals generated from the 5 Hz data, and accuracy of estimation tended to be better. Practical unidentifiability of the tractor's rear tire cornering stiffness, rear tire relaxation length, and implement tire relaxation length each emerged one SNR level lower for the 10 Hz data than for the 5 Hz data. The overall trends seen in the profile likelihood results, given in Figure 3.7, were the same as for the 5 Hz data. A compilation of confidence interval and estimation results for all three simulated cases is shown in Figure 3.9 and Figure 3.10.

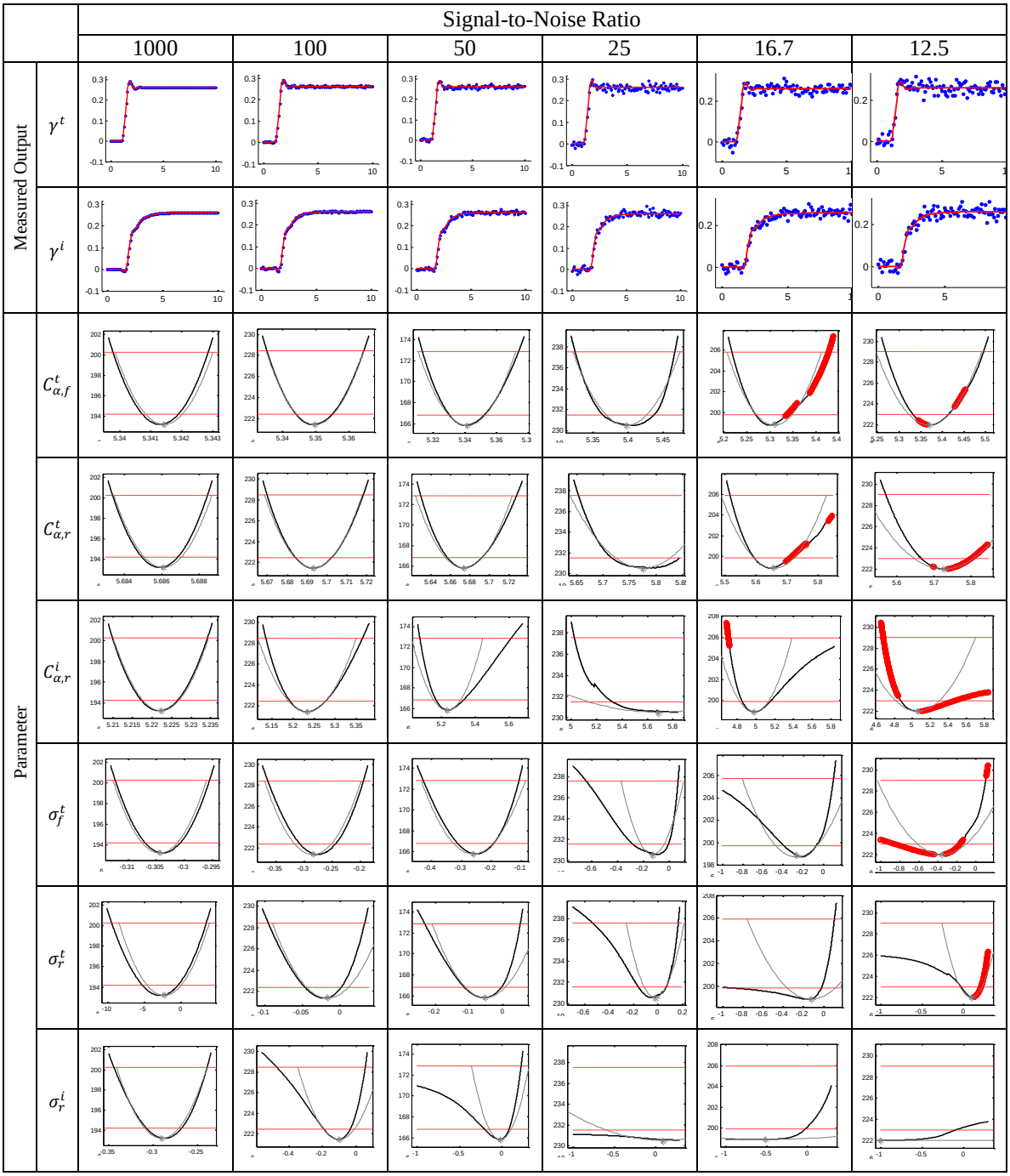


Figure 3.7 - Profile likelihood plots for each tire model parameter (in logarithmic space) when identified from simulated 10 Hz tractor yaw rate (rad/s) and implement yaw rate (rad/s) data for a 10 degree rate-limited step steer input. Simulations are grouped column-wise by the signal-to-noise ratio of the error applied to the measured outputs. Red dots on a profile likelihood plot indicate simulation points where at least one of the other five parameters was fit to one of its bounds. Discontinuities and smaller valleys encountered in the traversal of the profile likelihood are local minima.

3.3.1.3 Simulated Chirp Input Sampled at 5 Hz

Changing the steering signal used to excite the system from a rate-limited step input to a chirp input resulted in improvements in both accuracy and confidence of identification compared to both simulated cases 1 and 2. In many of the cases, the fitting operation led to optimum parameter values at or near the true values used to create the data. Likewise, the confidence intervals for the parameters identified using the chirp input were much narrower than their respective confidence intervals in simulated cases 1 and 2. Each of the tire cornering stiffnesses was identified relatively accurately compared to their true values, and each had finite upper and lower likelihood-based confidence bounds. Identification of the tire relaxation lengths was also slightly improved, and practical unidentifiabilities of each parameter emerged one SNR level lower than for the 10 Hz, rate-limited step input data in simulated case 2. The overall trends in the profile likelihood plots, shown in Figure 3.8, were the same as for the previous two simulated experiments. A compilation of results for all three simulated cases is shown in Figure 3.9 and Figure 3.10.

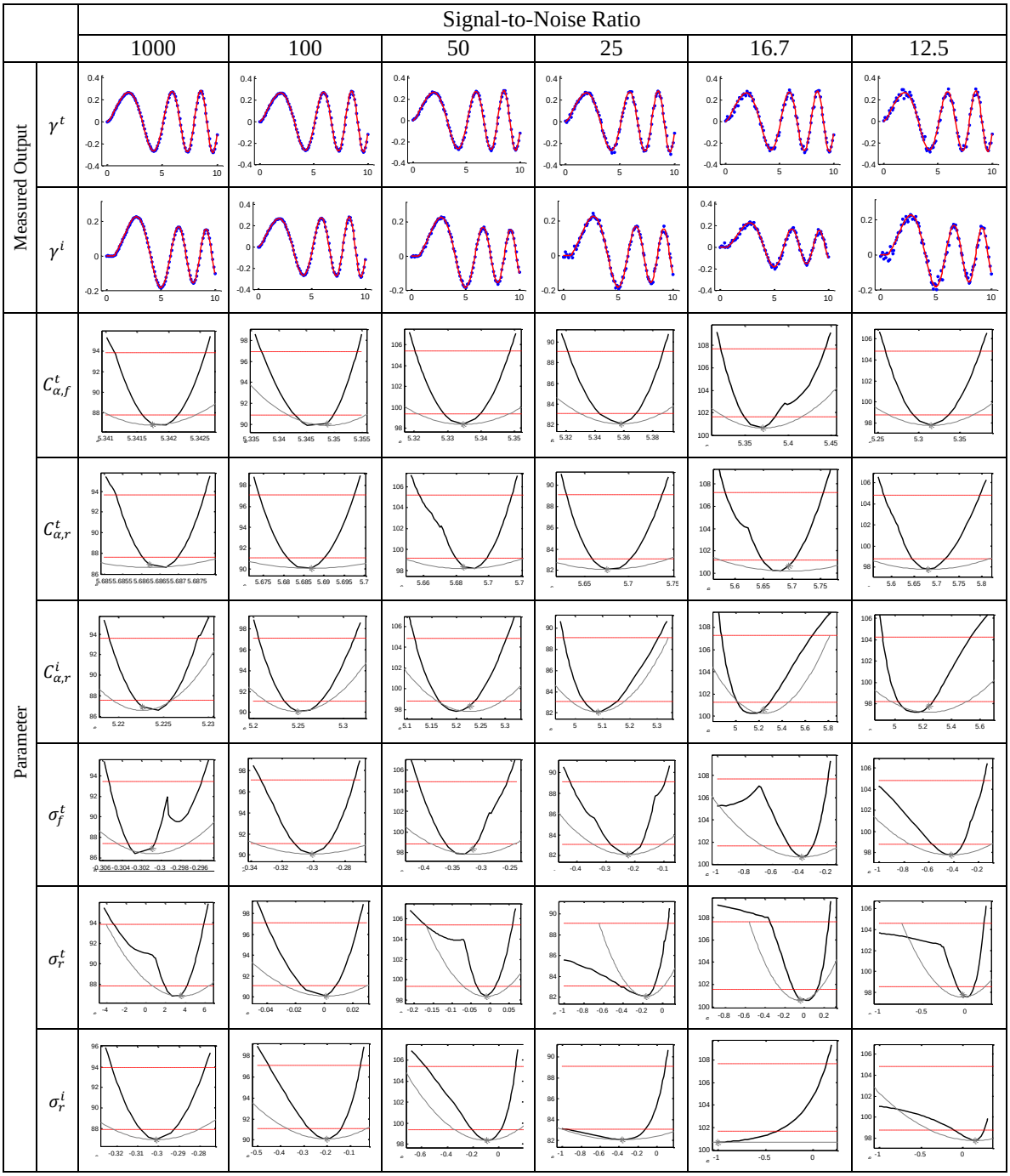


Figure 3.8 - Profile likelihood plots for each tire model parameter (in logarithmic space) when identified from simulated 5 Hz tractor yaw rate (rad/s) and implement yaw rate (rad/s) data for a 10 degree chirp steer input. Simulations are grouped column-wise by the signal-to-noise ratio of the error applied to the measured outputs. Red dots on a profile likelihood plot indicate simulation points where at least one of the other five parameters was fit to one of its bounds. Discontinuities and smaller valleys encountered in the traversal of the profile likelihood are local minima.

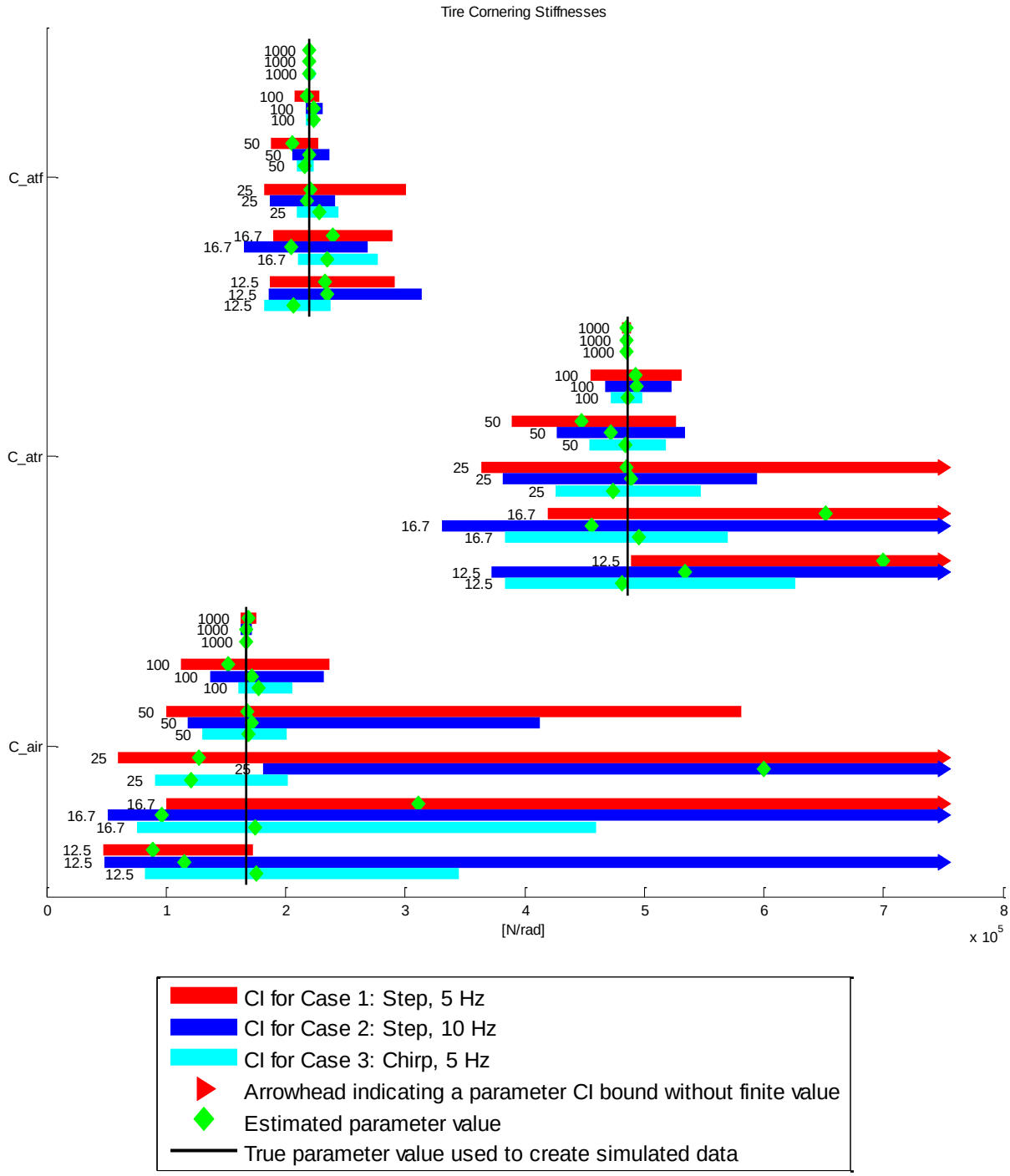


Figure 3.9 - Compilation of likelihood-based, 68% simultaneous confidence intervals (in normal space) for the tire cornering stiffness parameters in the three simulated data cases. The number to the left of each bar indicates the signal-to-noise ratio of the data it pertains to.

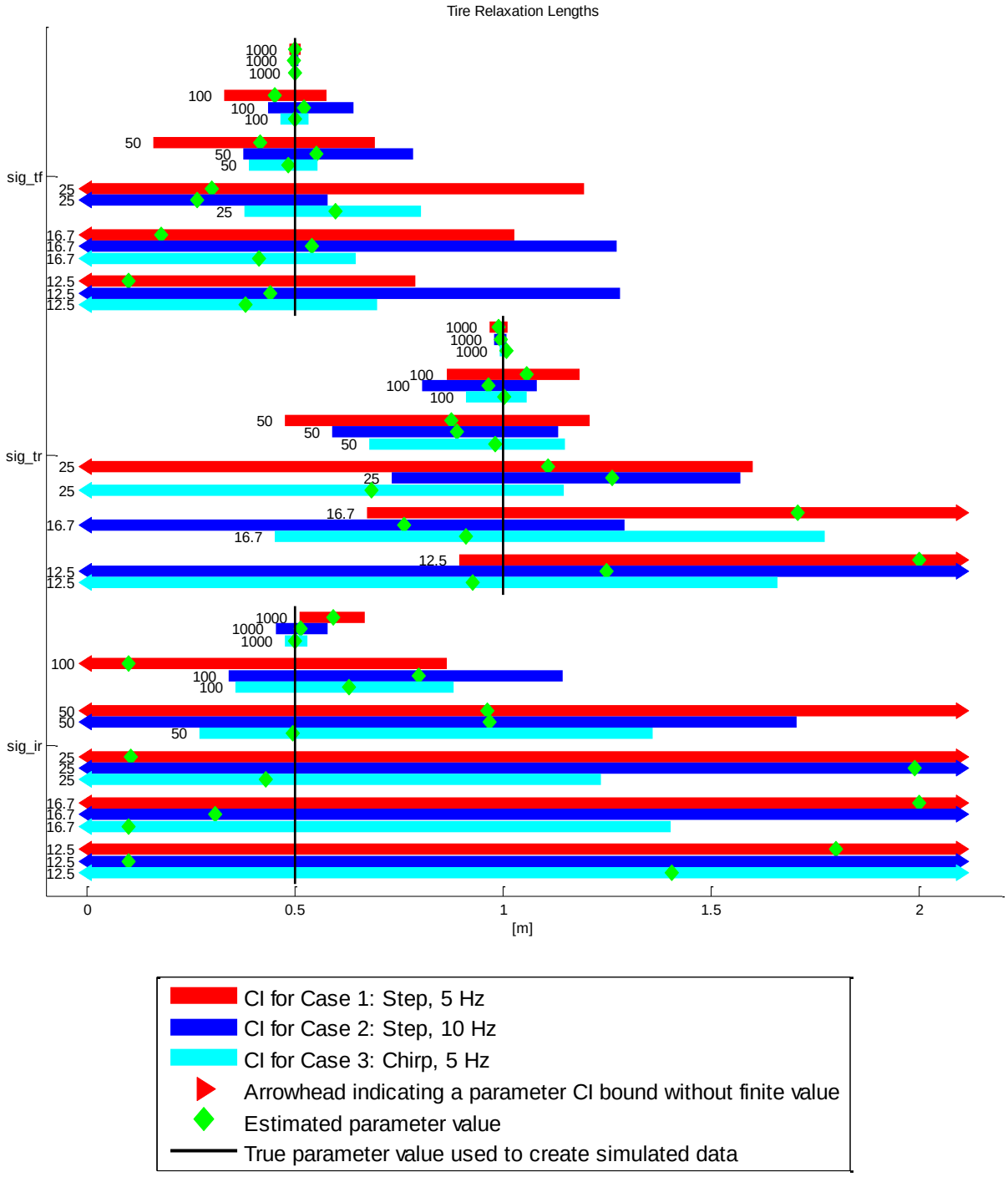


Figure 3.10 - Compilation of likelihood-based, 68% simultaneous confidence intervals (in normal space) for the tire relaxation length parameters in the three simulated data experiments. The number to the left of each bar indicates the signal-to-noise ratio of the data it pertains to.

3.3.2 Experimental Data Analysis

3.3.2.1 Step Steer Input

A 30-second long maneuver conducted at 4.5 m/s, consisting of a first step input to approximately -2 degrees and a second step input to approximately +2.5 degrees was examined. Time histories for the steer angle, tractor yaw rate, and implement yaw rate are shown in Figure 3.11. The experimental yaw rate data contained a significant amount of noise compared to the simulated datasets considered in this study. This noise was partially attributed to vehicle oscillations caused by the somewhat uneven ground surface, and it was larger at higher velocities (Karkee and Steward, 2011). However, it was still possible to visually observe the changes in the steady state yaw rate value during the step changes in the steering input. Karkee and Steward (2011) addressed the noise problem in their parameter identification study by identifying an error model as part of a prediction-error minimization (PEM) method. In the current study, the datasets were left unaltered to avoid elimination of any important dynamic information.

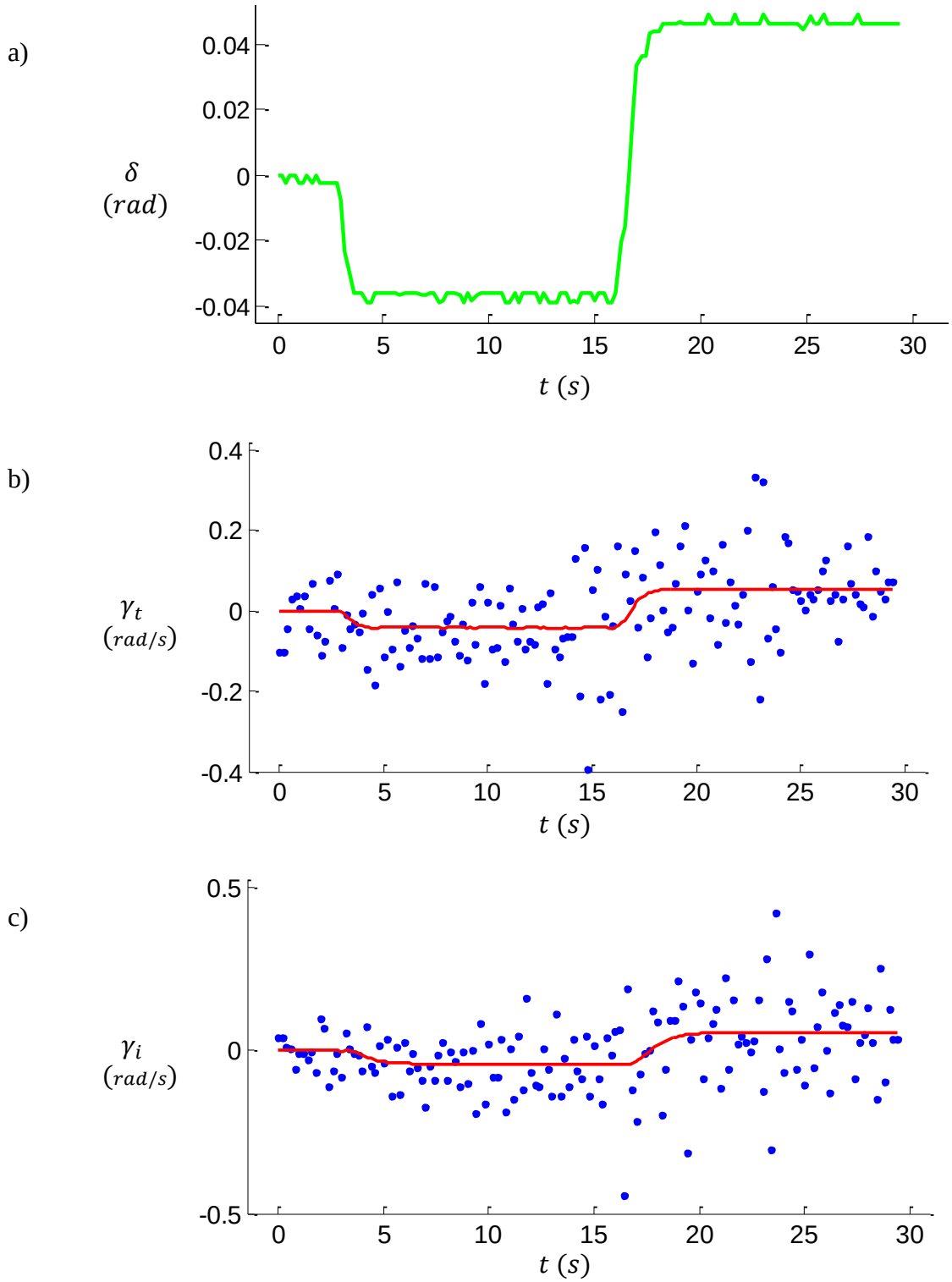


Figure 3.11 - Time histories of experimental data for step steering angle input: a) front wheel steering angle input, b) tractor yaw rate, and c) implement yaw rate. Blue dots are experimental data points, and the red lines are simulated outputs for the model after parameter estimation (in this study).

Quantile-quantile (Q-Q) plots of the residuals and histograms of weighted residuals for this data indicated that they were reasonably normally distributed. At the least, this suggested against the presence of substantial error in model structure.

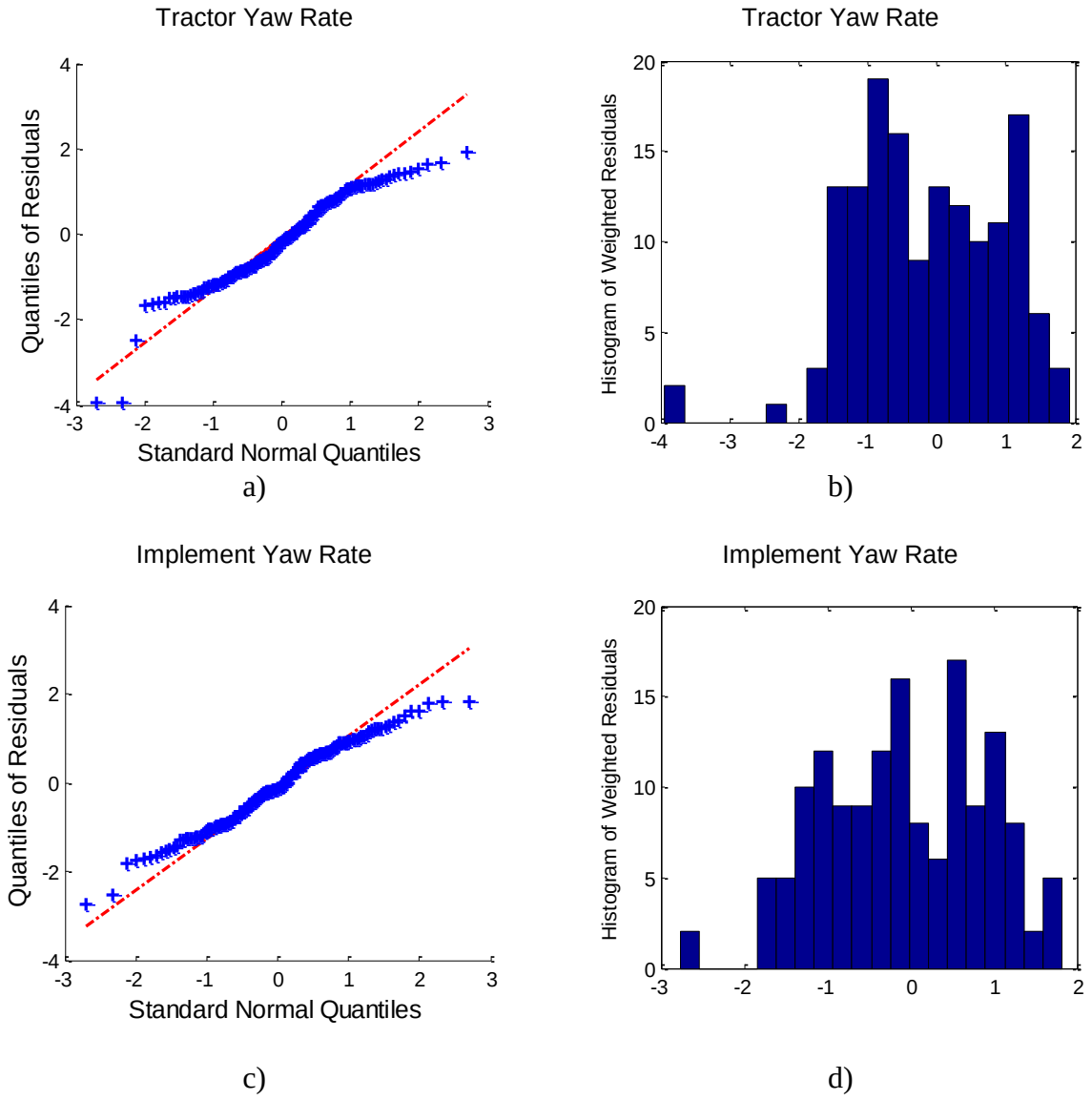


Figure 3.12 - Residuals for the experimental step-steer dataset with respect to the tractor-implement model with optimized parameters. A quantile-quantile (Q-Q) plot of the residuals and histogram of the weighted residuals for the tractor yaw rate are shown in a) and b), respectively. The same are shown for the implement yaw rate in c) and d), respectively.

The profile likelihood results indicated that only the tractor's front tire cornering stiffness was identifiable from the data, and all other parameters were practically unidentifiable.

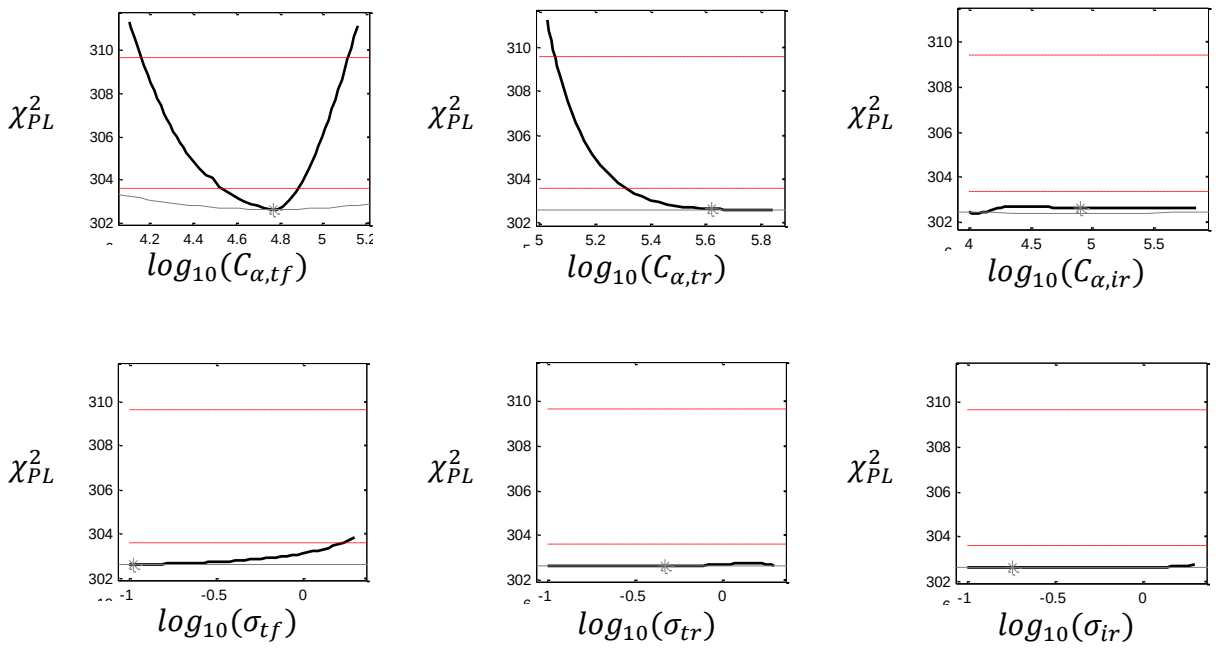


Figure 3.13 - Profile likelihoods for each of the six tire model parameters, plotted in logarithmic parameter space, for the experimental data in Figure 3.11. Black lines represent the profile likelihood; gray parabolas represent the quadratic approximation for asymptotic intervals. Gray asterisks at the valley of each curve indicate the estimated values of the parameters. The upper red dashed line of each plot represents the threshold for 68% simultaneous confidence intervals. The lower red dashed line represents the threshold for 68% pointwise confidence intervals.

It was possible to identify a lower bound for the tractor's rear tire cornering stiffness, but all other likelihood-based confidence bounds did not cross the upper χ^2 threshold within the range of parameter values allowed. In fact, the profile likelihoods for those unidentifiable parameters were very flat, showing little appreciable increase in the objective function as their values were varied across the allowed range. The results are summarized in Table 3.4. The identifiability of the parameters followed a trend similar to the results of the noisier data

sets in the simulated data analysis; that is, the parameters to which the output is more sensitive can be estimated with greater confidence.

Table 3.4 - Estimated values of each parameter and 68% simultaneous likelihood-based confidence intervals (all in normal parameter space) for the experimental data shown in Figure 3.11.

Parameter	$\hat{\theta}_i$	$CI_i^{-,PL}$	$CI_i^{+,PL}$
$C_{\alpha,tf}$	59300	14500	131100
$C_{\alpha,tr}$	421700	113600	$+\infty$
$C_{\alpha,ir}$	80700	0	$+\infty$
σ_{tf}	0.106	0	$+\infty$
σ_{tr}	0.471	0	$+\infty$
σ_{ir}	0.181	0	$+\infty$

The practical unidentifiability of the relaxation length parameters was expected based on the difficulty in visually observing any transient response in the yaw rate data. Practical unidentifiability of a parameter casts significant doubt on its estimated value, but it may be less important to have an accurate estimate of that parameter's value if its effect on the measured output is minimal.

3.3.2.2 Chirp Steer Input

Another 40-second long maneuver conducted at 4.5 m/s was considered which consisted of an approximate chirp steering angle input with maximum amplitudes between 5 and 8 degrees and frequencies between approximately 0.1 Hz and 0.5 Hz. The input was approximately sinusoidal but contained brief periods of constant steering angle at the wave peaks. Time histories for the steering angle, tractor yaw rate, and implement yaw rate are shown in Figure 3.14. The local spread of the data about the simulated trajectory was approximately the same as for the experimental step steer dataset, but the larger steering angle values in the chirp input resulted in larger yaw rate values. Therefore, the signal-to-noise ratio was larger than for the step steering dataset. The data were again left unaltered.

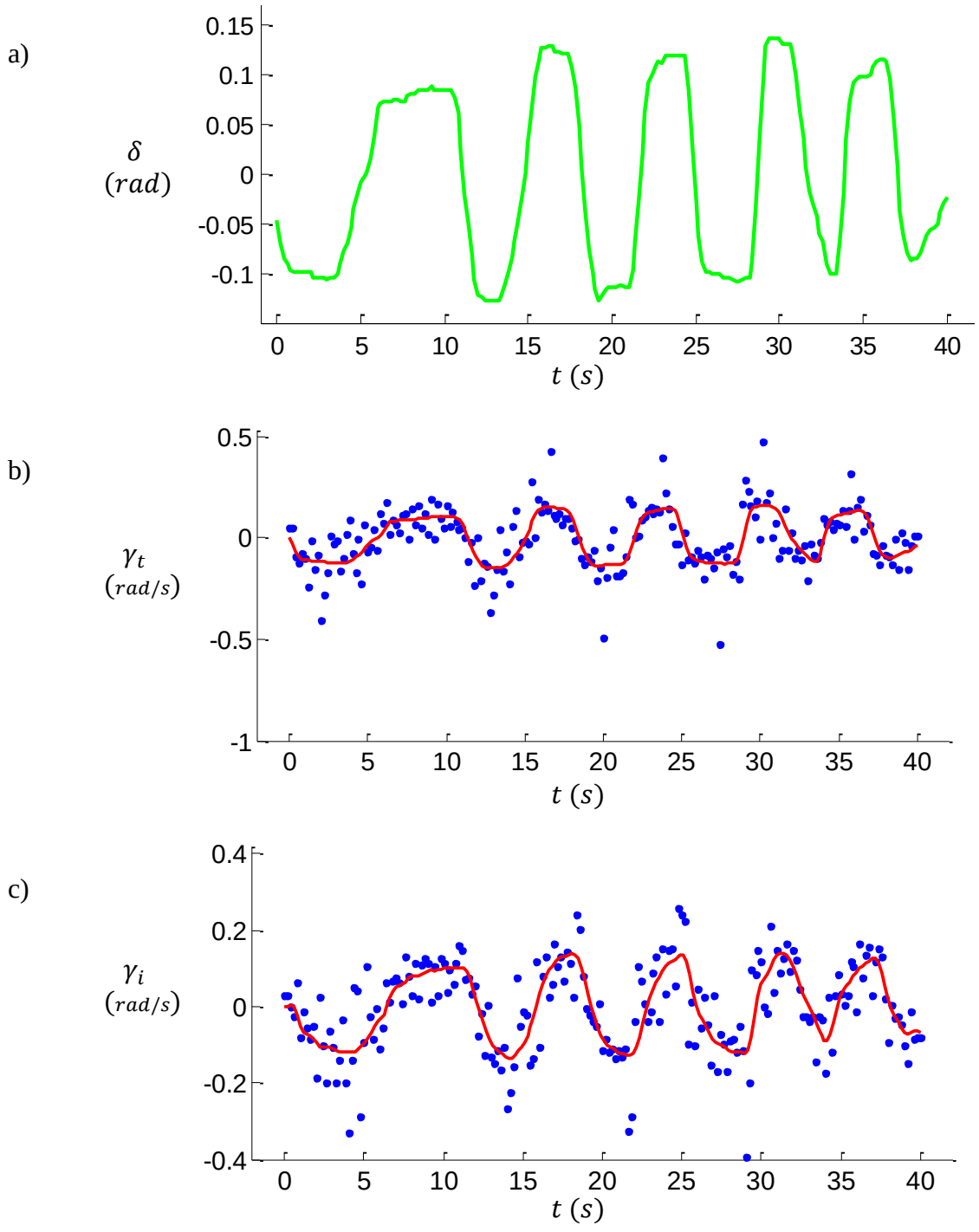


Figure 3.14 - Time histories (s) of experimental data for step steer input: a) front wheel steering input, b) tractor yaw rate, and c) implement yaw rate. Blue dots are experimental data points, and the red lines are simulated outputs for the model after parameter estimation (in this study).

Quantile-quantile (Q-Q) plots of the residuals and histograms of weighted residuals for this data, shown in Figure 3.15, indicated that they were reasonably normally distributed. At the least, this suggested against the presence of substantial error in model structure.

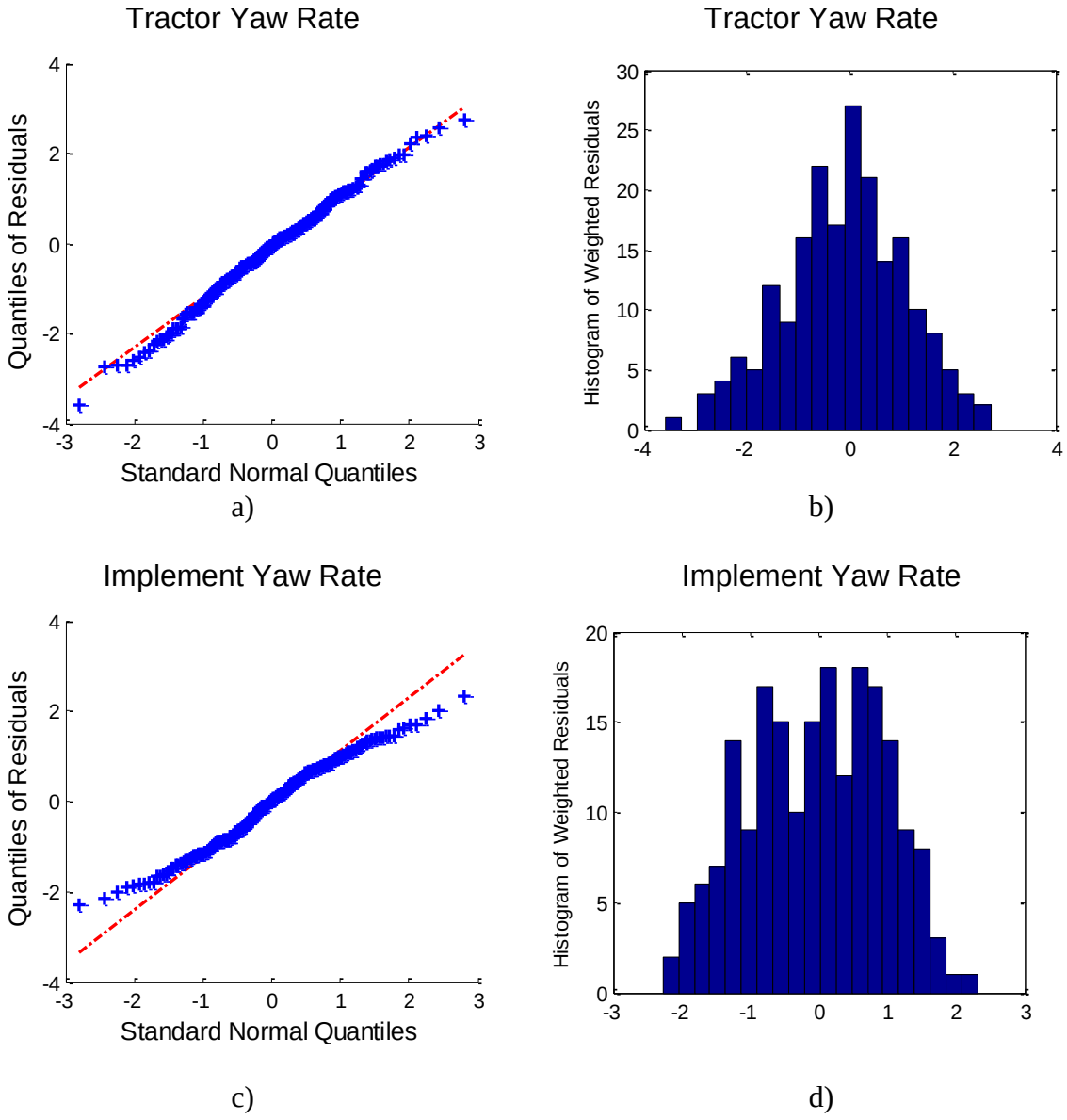


Figure 3.15 - Residuals for the experimental chirp-steer dataset with respect to the tractor-implement model with optimized parameters. A quantile-quantile (Q-Q) plot of the residuals and histogram of the weighted residuals for the tractor yaw rate are shown in a) and b), respectively. The same are shown for the implement yaw rate in c) and d), respectively.

The profile likelihood results again indicated that only the tractor's front tire cornering stiffness was practically identifiable. Four of the six parameters minimized the objective function at or near their upper or lower bounds, as shown in Table 3.5, automatically assigning them to be practically unidentifiable. Despite these extreme fits, a much greater amount of activity can be observed in the profile likelihood plots, shown in Figure 3.16, for the experimental chirp-input dataset than the experimental step-input dataset. This is possibly attributed to a greater amount of information available in the chirp input data since it contained a greater amount of transient response. Nonetheless, the results of the optimization and identifiability analysis again indicated that the model parameters were practically unidentifiable from the data available and encouraged the consideration of a reduced model compared to the dynamic model with tire relaxation length parameters considered here.

Table 3.5 - Estimated values of each parameter and 68% simultaneous likelihood-based confidence intervals (all in normal parameter space) for the experimental data shown in Figure 3.14.

Parameter	$\hat{\theta}_i$	$CI_i^{-,PL}$	$CI_i^{+,PL}$
$C_{\alpha,tf}$	72600	58600	91700
$C_{\alpha,tr}$	699800	558000	$+\infty$
$C_{\alpha,ir}$	166000	60800	$+\infty$
σ_{tf}	0.100	0	0.217
σ_{tr}	1.997	0	$+\infty$
σ_{ir}	1.940	0	$+\infty$

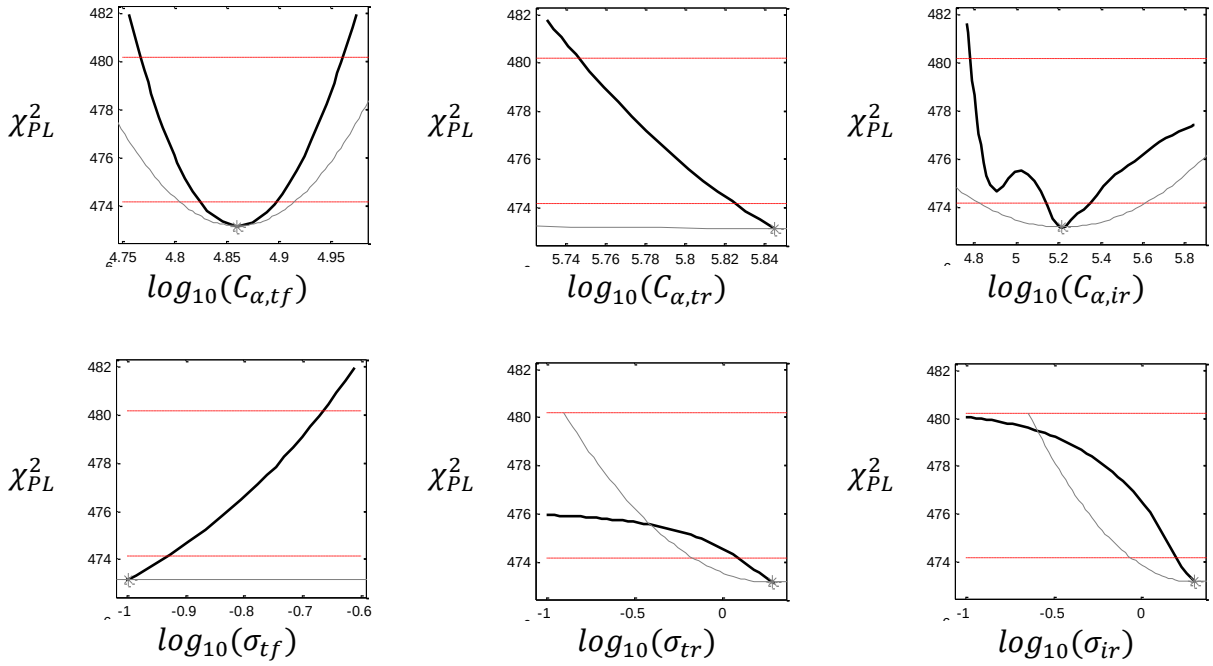


Figure 3.16 - Profile likelihoods for each of the six tire model parameters, plotted in logarithmic parameter space, for the experimental data in Figure 3.14. Black lines represent the profile likelihood; gray parabolas represent the quadratic approximation for asymptotic intervals. Gray asterisks at the valley of each curve indicate the estimated values of the parameters. The upper red dashed line of each plot represents the threshold for 68% simultaneous confidence intervals. The lower red dashed line represents the threshold for 68% pointwise confidence intervals.

3.4 Conclusions

The profile likelihood approach to identifiability analysis proposed by Raue et al. (2009) provided a numerical means by which to examine a model and was able to incorporate actual data. The approach showed that even structurally identifiable parameters become practically unidentifiable due to the presence of noise and other properties related to data collection. The shape of the likelihood for each parameter provided additional information about its identifiability and lent insight into opportunities for model reduction.

The simulated data analysis showed the challenge of parameter identification even for data generated from a model with constant tire model parameters and a Gaussian noise

model. Although the tractor-implement model was structurally identifiable (at least locally), the issue of greater concern was practical identifiability due to the properties of the data. Although the tractor and implement yaw rates had non-zero sensitivities to each of the six tire model parameters (Karkee and Steward, 2010b), several of the parameters, especially the relaxation lengths and the implement cornering stiffness, were more difficult to identify accurately as measurement noise increased and were often practically unidentifiable in terms of their likelihood-based confidence intervals. Also, the parameter values that minimized the objective function often pertained to estimates that were far from the true values. Overall, it showed that the model structure and the experimental design can have an impact on the identifiability of parameters.

The experimental data analysis showed that, although it was possible to identify a set of tire model parameters for the tractor-implement model that minimized the error between the simulated and experimental yaw rates, all but the tractor's front tire cornering stiffness were considered practically unidentifiable from the data available. In both datasets considered, at least two of the six parameters were fit to an upper or lower bound. Efforts to decrease the amount of noise in the data, increase the collection frequency, measure additional outputs, or perform maneuvers which more-persistently excite the system would likely improve the accuracy and confidence of identification in future experiments.

It should be noted that the confidence interval values and the identifiability conclusions for every dataset in this study are directly related to the confidence level chosen as well as the bounds chosen for each parameter. Modification of either the confidence level or the parameter bounds could change the conclusions made, but those values need to be selected based on the purpose of the model identification in any given study. The confidence

level for the results obtained in this study could be decreased without rerunning the analysis by simply calculating a different χ^2 threshold and determining where it intersects the profile likelihood plots.

Although this study did not provide any direct improvement upon off-road vehicle modeling or the identification of tire model parameters themselves, it demonstrated the concept of identifiability in that context and lent insight into the ongoing challenge of model parameter identification from experimental data. Specifically, it challenges researchers to design experiments that maximize the amount of information available in the measured data, to minimize the presence of noise in the data, and to choose levels of model fidelity that can be reliably identified. Identification from experimental data is likely to increase in use as model-based design of off-road vehicle systems becomes more common and as sensing capabilities improve.

In summary, the following conclusions can be drawn from this work:

- The tractor-implement model with six free tire model parameters and measured tractor and implement yaw rates is structurally identifiable, at least locally. However, aspects related to experimental data collection, such as measurement noise, sampling rate, and the system input excitation can affect the practical identifiability of the model. As data quality and quantity decreased, the practical unidentifiability of model parameters tended to emerge in the order of least to greatest sensitivity.
- The numerical, profile likelihood approach for testing structural and practical identifiability lent additional insight into the identification of the tractor-implement model. The shape of the likelihood indicated opportunities for model reduction that might lead to more reliable identification of parameters. In the simulated data

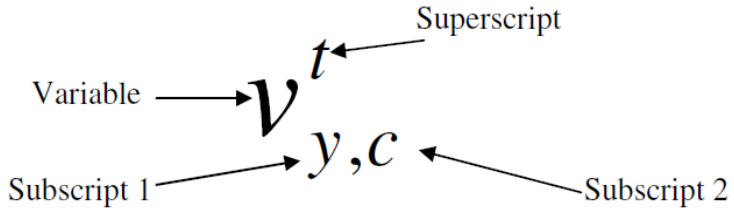
analysis, the size of the likelihood-based confidence intervals tended to indicate the accuracy of the identification as well.

- The unprocessed data collected from two representative field experiments appeared to be of insufficient quality and quantity for reliable identification of the tractor-implement model parameters. Efforts to increase the number of measured outputs, increase sampling rates, and utilize input signals that maximize the richness of the data would likely improve the results of identification. Sensor noise may be mitigated with the use of on- or off-line signal filters, but this will not necessarily improve the accuracy of identification, especially if it removes important dynamic information. Finally, depending on the application, a reduced model with fewer free parameters may be more reliably identifiable from the data available.

Acknowledgements

We acknowledge Dr. Manoj Karkee for his previous tractor-implement modeling and experimental data collection.

Notation and List of Variables



Variable: the variable itself

Superscript: denotes whether the variable is related to the tractor or the implement
t – tractor, i – implement

Subscript 1: specifies the coordinate axis the variable corresponds to
x – x-axis, y – y-axis, z – z-axis

Subscript 2: specifies the location the variable corresponds to
f – front, r – rear, c – center of gravity

Tractor-implement model

α	tire lateral slip angle
α_0	steady-state tire lateral slip angle
γ	yaw rate
δ	wheel steer angle
σ	tire relaxation length
ψ	heading angle
a	distance between front axle and CG of tractor
b	distance between rear axle and CG of tractor
c	distance between hitch point and CG of tractor
C_α	tire cornering stiffness
d	distance between hitch point and CG of implement
e	distance between rear axle and CG of implement
F	force
I	yaw moment of inertia
m	mass
u	longitudinal velocity
v	lateral velocity
x	position of a CG in the x-axis of the world coordinate system
y	position of a CG in the y-axis of the world coordinate system

Identification

$\boldsymbol{\theta}$	vector of model parameters to be estimated
θ_i	parameter of index i in θ

$\boldsymbol{\theta}^*$	true value of θ
$\widehat{\boldsymbol{\theta}}$	estimate of θ
σ	measurement error
χ_k^2	chi-square distribution with k degrees of freedom
CI	confidence interval bound
df	degrees of freedom
$\boldsymbol{e}_y$	output error
$M(\boldsymbol{\theta})$	model with structure M and parameter vector θ
PL	profile likelihood abbreviation
$\mathbb{P}$	prior feasible set for θ
$\mathbb{R}^n$	set of real numbers in n-dimensional space
SNR	signal-to-noise ratio
t	time
$\boldsymbol{u}$	vector of controlled inputs to a model or system
$\boldsymbol{x}$	model state vector
$\boldsymbol{y}$	system output vector
$\boldsymbol{y_m}$	model output vector

References

Ahmed, O. B., and J. F. Goupillon. (1997). Predicting the ride vibration of an agricultural tractor. *Journal of Terramechanics*, 34(1): 1-11.

Alasty, A., and A. Ramezani. (2002). Genetic algorithm based parameter identification of a nonlinear full vehicle ride model. *SAE Paper No. 2002-01-1583*.

Arikan, K. B. (2008). Identification of handling models for road vehicles. *PhD thesis*. Middle East Technical University.

Balsa-Canto, E., and J. R. Banga. (2010). Computational procedures for model identification. In: Systems Biology for Signaling Networks. S. Choi (ed.), 111-137. New York: *Springer.*

Bernard, J. E., and C. L. Clover. (1994). Validation of computer simulations of vehicle dynamics. *SAE International Congress & Exposition*. Feb. 28 – Mar. 3, Detroit, MI. SAE Paper No. 940231.

Bevly, D. M., J. C. Gerdes, and B. W. Parkinson. (2002). A new yaw dynamic model for improved high speed control of a farm tractor. *Journal of Dynamic Systems, Measurement, and Control*, 124: 659-667.

Book, R. S., and C. E. Goering. (2000). A new traction model for crawler tractors. *Transactions of the ASABE*, 43(1): 39-46.

Crolla, D. A. and A. S. A. El-Razaz. (1987). A review of the combined lateral and longitudinal force generation of tyres on deformable surfaces. *Journal of Terramechanics*, 24(3): 199-225.

Dieter, G. E., and L. C. Schmidt. (2009). Engineering Design. Fourth Edition. New York: *McGraw-Hill*.

Fales, R., E. Spencer, K. Chipperfield, F. Wagner, and A. Kelkar. (2005). Modeling and control of a wheel loader with human-in-the-loop assessment using virtual reality. *Journal of Dynamic Systems, Measurement, and Control*, 127: 415-423.

Gillespie, T. D. (1992). Fundamentals of Vehicle Dynamics. Warrendale, PA: *Society of Automotive Engineers, Inc.*

Goldberg, D. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning. *Addison Wesley*.

Hengl, S., C. Kreutz, J. Timmer, and T. Maiwald. (2007). Data-based identifiability analysis of non-linear dynamical models. *Bioinformatics*. 23: 2612-2618.

Hindmarsh, A. C., P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward. (2005). SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software*, 31: 363-396.

Hummel, J., E. Baack, and J. Bernard. (2005). Operator in the loop simulation for farm and construction vehicles. *DSC 2005 North America – Orlando,* 421-432.

Jang, J. H., and C. S. Han. (1997). The state sensitivity analysis of the front wheel steering vehicle: In the time domain. *KSME International Journal*, 11(6): 595-604.

Karimi, D., and D. D. Mann. (2006). A study of tractor yaw dynamics for application in a tractor driving simulator. *ASABE/CSBE North Central Intersectional Meeting*. Paper No. MBSK 06-113.

Karkee, M. (2009). Modeling, identification and analysis of tractor and single axle towed implement system. *PhD dissertation,* Iowa State University.

Karkee, M., B. L. Steward, A. G. Kelkar, and Z. T. Kemp II. (2009). Modeling and real-time simulation architectures for virtual prototyping of off-road vehicles. *Virtual Reality*, 15: 83-96.

Karkee, M., and B. L. Steward. (2010a). Study of the open and closed loop characteristics of a tractor and a single axle towed implement system. *Journal of Terramechanics*, 47(6): 379-393.

Karkee, M., and B. L. Steward. (2010b). Local and global sensitivity analysis of a tractor and single axle grain cart dynamic model. *Biosystems Engineering*, 106(4): 352-366.

Karkee, M., and B. L. Steward. (2011). Parameter estimation and validation of a tractor and single axle towed implement dynamic system model. *Computers and Electronics in Agriculture*. In press.

Kiencke, U., and L. Nielsen. (2005). Automotive Control Systems for Engine, Driveline, and Vehicle. Second Edition. Berlin: *Springer*.

Koolen, A. J., and H. Kuipers. (1983). Agricultural Soil Mechanics. Berlin: *Springer-Verlag*.

Lennon, T. (2008). Model-based design for mechatronic systems. *EngineerIT*, April: 16-20.

Ljung, L. (1999). System Identification: Theory for the User. Second Edition. Englewood Cliffs, NJ: *Prentice Hall*.

Maiwald, T., and J. Timmer. (2008). Dynamical modeling and multi-experiment fitting with PottersWheel. *Bioinformatics*, 24: 2037-2043.

Meeker, W. Q., and L. A. Escobar. (1998). Statistical Methods for Reliability Data. *John Wiley & Sons, Inc.*

Metz, L. D. (1993). Dynamics of four-wheel-steer off-highway vehicles. *SAE Paper No. 930765*.

Nocedal, J, and S. J. Wright. (1999). Numerical Optimization. New York: *Springer*.

Pacejka, H. B. (2006). Tire and Vehicle Dynamics, Second Edition. Warrendale, PA: *Society of Automotive Engineers, Inc.*

Prabhu, S. (2007). Model-based design for off-highway machine systems development. *SAE 2007 Commercial Vehicle Engineering Congress & Exhibition.* SAE Paper No. 2007-01-4248.

Previati, G., M. Gobbi, and G. Mastinu. (2007). Farm tractor models for research and development purposes. *Vehicle System Dynamics*, 45: 37-60.

Radhakrishnan, R., and D. A. McAdams. (2005). A methodology for model selection in engineering design. *ASME Journal of Mechanical Design*, 127: 378-387.

Raue, A., C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15): 1923-1929.

Raue, A., C. Kreutz, T. Maiwald, U. Klingmüller, and J. Timmer. (2011). Addressing parameter identifiability by model-based experimentation. *IET Systems Biology*, 5(2): 120-130.

Serban, R., and J. S. Freeman. (2001). Identification and identifiability of unknown parameters in multibody dynamic systems. *Multibody System Dynamics*, 5: 335-350.

Venkataraman, P. (2009). Applied optimization with MATLAB programming. Second Edition. *John Wiley & Sons, Inc.*

Walter, E., and L. Pronzato. (1997). Identification of Parametric Models from Experimental Data. *Springer*.

Wong, J. Y. (1989). Terramechanics and Off-Road Vehicles. Amsterdam, The Netherlands: *Elsevier*.

Wong, J. Y. (2008). Theory of Ground Vehicles. Fourth Edition. Hoboken, NJ: *John Wiley & Sons, Inc.*

Wymore, A. W. (1993). Model-Based Systems Engineering. Boca Raton, FL: *CRC Press.*

Åström, K. J., H. Elmqvist, and S. E. Mattsson. (1998). Evolution of continuous-time modeling and simulation. *In: Proceedings of the 12th European Simulation Multiconference, ESM'98.* June 16-19, Manchester, UK.

CHAPTER 4. GENERAL CONCLUSIONS

4.1 Conclusions

The overall objective of this research was to investigate methods that could be used to assess the process of identifying parametric models from experimental data. After reviewing the current trends in model-based design and parameter identification from experimental data, especially in vehicle dynamics applications, identifiability was recognized as a model property that could lend insight into the investigation. Nonetheless, many mechanical system identification studies do not appear to consider identifiability as part of their analysis. A number of methods for testing structural identifiability and practical identifiability of models were found in the literature; particularly, a numerical approach which could test for local structural and practical identifiability of arbitrary models was recognized. This approach considered the statistical aspects of parameter identification from actual data through principles of maximum likelihood estimation and, specifically, used the profile likelihood of each parameter to determine likelihood-based confidence intervals and assess their identifiability. An implementation of the approach in a third-party mathematical modeling toolbox for MATLAB was demonstrated on two simple mechanical system examples.

The additional objectives of this research were to investigate the identifiability of a tractor and single axle towed implement model from measured data and to evaluate the effects of experimental factors such as measurement noise, data sampling rates, and input excitation type on the identification process. The identifiability of the six tire force model parameters from tractor yaw rate and implement yaw rate data was investigated using simulated data with fixed parameter values, which enabled a more controlled assessment of

the three previously-mentioned experimental factors and eliminated any possibility of error in model structure. The analysis showed that the six unknown model parameters were at least locally, structurally identifiable, but measurement noise, sampling rate, and input excitation type had an impact on both the accuracy and confidence in what could be considered an ideal scenario for identification. The trends in the profile likelihood plots indicated opportunities for model reduction as well. Identifiability analysis of the parameters from actual experimental datasets suggested that the tractor yaw rate and implement yaw rate data available did not contain enough information for several of the parameters to be considered practically identifiable to the level of confidence specified.

Overall, it has become increasingly clear that identification of models from experimental data needs to be carefully planned and conducted in order to increase the chances of success in estimation. The approach should be multifaceted, considering model structure, experimental design, sensor configuration, signal processing, optimization methods, and validation. Ultimately, the objectives of any modeling effort will determine what level of accuracy/realism is needed for identification to be considered "successful" and for a model to be considered "useful".

4.2 Recommendations for Future Work

There are several areas where this work could be improved or expanded. Some potential areas are described as follows:

- In this work, one experimental factor that was not considered in the identifiability analysis of the tractor-implement model was the impact of different sensor configurations. An additional simulated data analysis case could consider the

selection and placement of sensors on the tractor and implement to optimize the types of information collected in an experiment.

- This study only considered the identifiability of the six tire model parameters in the tractor-implement model and assumed that the remaining model parameters (masses, inertias, and geometric dimensions) were known with certainty. A more complete study would include the identification of these additional parameters from experimental data and assess their impact on the identifiability of the model.
- Before one can properly consider local identifiability of a model with respect to a location in the parameter space, it must be assumed that the global optimum of the model parameters has been estimated. Reaching the global optimum is a challenge of its own, especially as parameter set dimensions increase and as an objective function landscape contains more nonlinearities and local minima. Further insight into the convergence properties of global optimization techniques like genetic algorithms as well as the development of optimal objective functions for problems in off-road vehicle model identification should be considered.
- Numerical identifiability approaches could be implemented for handling arbitrary models in common modeling packages like MATLAB/Simulink. Although a number of MATLAB/Simulink toolboxes (e.g., System Identification, Simulink Design Optimization, Optimization, and Statistics, among others) contain useful functions for model identification, they do not appear to have any functions built-in for specifically examining structural nor practical identifiability of models. Although this study relied heavily on a numerical identifiability method, analytical methods should be considered when it is feasible.

- The literature review conducted during this study has raised further awareness to the academic areas of system identification, estimation theory, and optimal experimental design. It is suggested that educational efforts in vehicle dynamics, modeling, and simulation be supplemented with theory from these previously mentioned areas in future model identification projects to improve the chances for success and to better-design experiments, which can be expensive, time-consuming, and highly dependent upon environmental conditions.

APPENDIX. POTTERSWHEEL FILE FOR THE TRACTOR-IMPLEMENT MODEL

```
% PottersWheel model definition file

function m = TracImpDynamicwRL()

m = pwGetEmptyModel();

%% Meta information

% Dynamic tractor-implement model
m.t = [0:0.2:10];                                   % 5 Hz data collection

%% Dynamic variables
% m = pwAddX(m, ID, startValue, type, minValue, maxValue)

m = pwAddX(m, 'v_tc', 0);                           % [m/s]; tractor CG
lateral velocity
m = pwAddX(m, 'gam_t', 0);                          % [rad/s]; tractor yaw
rate
m = pwAddX(m, 'gam_i', 0);                          % [rad/s]; implement
yaw rate
m = pwAddX(m, 'alp_tf', 0);                         % [rad]; tractor front
tire slip angle
m = pwAddX(m, 'alp_tr', 0);                         % [rad]; tractor rear
tire slip angle
m = pwAddX(m, 'alp_ir', 0);                         % [rad]; implement
tire slip angle
m = pwAddX(m, 'x_tc', 0);                           % [m]; tractor CG
trajectory in world
m = pwAddX(m, 'y_tc', 0);                           % [m]; tractor CG
trajectory in world
m = pwAddX(m, 'psi_t', 0);                          % [rad]; tractor
heading angle in world
m = pwAddX(m, 'psi_i', 0);                          % [rad]; implement
heading angle in world

%% Dynamic parameters
% m = pwAddK(m, ID, value, fitSetting, minValue, maxValue, unit, name,
description)

% p130
m = pwAddK(m, 'C_atf', 220000, [], 10000, 700000);  % [N/rad]; tractor
front tire cornering stiffness
m = pwAddK(m, 'C_atr', 486000, [], 10000, 700000);  % [N/rad]; tractor
rear tire cornering stiffness
m = pwAddK(m, 'C_air', 167000, [], 10000, 700000);  % [N/rad]; implement
tire cornering stiffness
m = pwAddK(m, 'sig_tf', 0.5, [], 0.1, 2.0);         % [m]; tractor front
tire relaxation length
```

```
m = pwAddK(m, 'sig_tr', 1.0, [], 0.1, 2.0);          % [m]; tractor rear
tire relaxation length
m = pwAddK(m, 'sig_ir', 0.5, [], 0.1, 2.0);          % [m]; implement tire
relaxation length

%% Constant (fixed) parameters

% p130
m = pwAddK(m, 'm_t', 9391, 'fix');                   % [kg]; tractor mass
m = pwAddK(m, 'm_i', 2127, 'fix');                   % [kg]; implement mass
m = pwAddK(m, 'I_tz', 35709, 'fix');                 % [kg-m^2]; tractor
yaw moment of inertia
m = pwAddK(m, 'I_iz', 6402, 'fix');                  % [kg-m^2]; implement
yaw moment of inertia
m = pwAddK(m, 'u_tc', 4.5, 'fix');                   % [m/s]; tractor
longitudinal velocity
m = pwAddK(m, 'a', 1.7, 'fix');                      % [m]; distance
between front axle and CG of tractor
m = pwAddK(m, 'b', 1.2, 'fix');                      % [m]; distance
between rear axle and CG of tractor
m = pwAddK(m, 'c', 2.1, 'fix');                      % [m]; distance
between tractor CG and hitch point
m = pwAddK(m, 'd', 3.62, 'fix');                     % [m]; distance
between hitch point and implement CG
m = pwAddK(m, 'e', 0.1, 'fix');                      % [m]; distance
between implement CG and implement axle

%% Driving input

% Ramp
m = pwAddU(m, 'del', 'linear', [0 1 1.5], [0 0 0.17452]);  % [rad]; ramp
input to front wheel angle

% Chirp
% m = pwAddU(m, 'del', 'linear', [0:0.1:10],
[0,0.0111778960931067,0.0227448762788643,0.0346410333652654,0.046794516355
9924,0.0591209200963829,0.0715228376915582,0.0838896160672813,0.0960973576
325432,0.108009212795761,0.119476008834540,0.130337260052100,0.14042260198
6874,0.149553688384379,0.157546583405930,0.164214672869473,0.1693721069614
96,0.172837772648374,0.174439776850454,0.174020401324205,0.171441467255787
,0.166590022086360,0.159384233528051,0.149779346755508,0.137773531259491,0
.123413414948844,0.106799076146579,0.0880882407358103,0.0674994136696496,0
.0453136633304624,0.0218747758786817,-0.00241249313492610,-
0.0270863189917560,-0.0516340700304305,-0.0755010863562033,-
0.0981020561425106,-0.118834979337074,-0.137097600015138,-
0.152306064481649,-0.163915425428713,-0.171441467255787,-
0.174483179714332,-0.172745063408684,-0.166058319751008,-
0.154399869234677,-0.137908065670923,-0.116893941005801,-
0.0918468359616914,-0.0634333555713249,-0.0324887434952374,-
4.27482376382620e-
17,0.0329196227850760,0.0650647311188034,0.0951800379460785,0.122009779038
965,0.144352792003766,0.161121213152131,0.171400234580817,0.17450591598395
0,0.170037703428543,0.157922112266475,0.138444023440735,0.112262256453567,
0.0804065441731334,0.0442537577970197,0.00548221166464041,-
```

```
0.0339959053295080,-0.0721225166557922,-0.106799076146579,-
0.136004863457722,-0.157922112266475,-0.171060634823028,-
0.174373646072444,-0.167355991317692,-0.150116059953858,-
0.123413414948844,-0.0886556353884071,-0.0478500583535319,-
0.00350895624360309,0.0414899317953503,0.0840818779665467,0.12122322635399
7,0.150116059953858,0.168433539395435,0.174527964218666,0.167602866011742,
0.147831238905102,0.116404510817839,0.0755010863562026,0.0281691133891860,
-0.0218747758786819,-0.0705211180823961,-0.113600029657880,-
0.147245386376670,-0.168259917990565,-0.174446803736655,-
0.164872798225465,-0.140030855354542,-0.101876837333062,-
0.0537250074001185,-1.28244712914786e-16;]);

%% ODEs
% m = pwAddODE(m, leftHandSide, rightHandSide)

m = pwAddODE(m, 'v_tc', '- gam_t*u_tc - (C_atf*alp_tf*(I_iz*m_i*c^2 +
I_iz*a*m_i*c + I_tz*m_i*d^2 + I_iz*I_tz))/(I_iz*m_i*m_t*c^2 +
I_tz*m_i*m_t*d^2 + I_iz*I_tz*m_i + I_iz*I_tz*m_t) -
(C_atr*alp_tr*(I_iz*m_i*c^2 - I_iz*b*m_i*c + I_tz*m_i*d^2 +
I_iz*I_tz))/(I_iz*m_i*m_t*c^2 + I_tz*m_i*m_t*d^2 + I_iz*I_tz*m_i +
I_iz*I_tz*m_t) - (C_air*I_tz*alp_ir*(I_iz - d*e*m_i))/(I_iz*m_i*m_t*c^2 +
I_tz*m_i*m_t*d^2 + I_iz*I_tz*m_i + I_iz*I_tz*m_t)');
m = pwAddODE(m, 'gam_t', '(C_atr*alp_tr*(b*m_i*m_t*d^2 + I_iz*b*m_i +
I_iz*b*m_t - I_iz*c*m_i))/(I_iz*m_i*m_t*c^2 + I_tz*m_i*m_t*d^2 +
I_iz*I_tz*m_i + I_iz*I_tz*m_t) - (C_atf*alp_tf*(a*m_i*m_t*d^2 + I_iz*a*m_i
+ I_iz*a*m_t + I_iz*c*m_i))/(I_iz*m_i*m_t*c^2 + I_tz*m_i*m_t*d^2 +
I_iz*I_tz*m_i + I_iz*I_tz*m_t) + (C_air*alp_ir*c*m_t*(I_iz -
d*e*m_i))/(I_iz*m_i*m_t*c^2 + I_tz*m_i*m_t*d^2 + I_iz*I_tz*m_i +
I_iz*I_tz*m_t)');
m = pwAddODE(m, 'gam_i', '(C_air*alp_ir*(e*m_i*m_t*c^2 + I_tz*d*m_t +
I_tz*e*m_i + I_tz*e*m_t))/(I_iz*m_i*m_t*c^2 + I_tz*m_i*m_t*d^2 +
I_iz*I_tz*m_i + I_iz*I_tz*m_t) - (C_atf*alp_tf*d*m_i*(I_tz -
a*c*m_t))/(I_iz*m_i*m_t*c^2 + I_tz*m_i*m_t*d^2 + I_iz*I_tz*m_i +
I_iz*I_tz*m_t) - (C_atr*alp_tr*d*m_i*(I_tz + b*c*m_t))/(I_iz*m_i*m_t*c^2 +
I_tz*m_i*m_t*d^2 + I_iz*I_tz*m_i + I_iz*I_tz*m_t)');
m = pwAddODE(m, 'alp_tf', 'v_tc/sig_tf - (alp_tf*u_tc)/sig_tf -
(del*u_tc)/sig_tf + (a*gam_t)/sig_tf');
m = pwAddODE(m, 'alp_tr', 'v_tc/sig_tr - (alp_tr*u_tc)/sig_tr -
(b*gam_t)/sig_tr');
m = pwAddODE(m, 'alp_ir', 'v_tc/sig_ir - (alp_ir*u_tc)/sig_ir -
(psi_i*u_tc)/sig_ir + (psi_t*u_tc)/sig_ir - (gam_i*(d + e))/sig_ir -
(c*gam_t)/sig_ir');
% m = pwAddODE(m, 'y_tc', 'v_tc + psi_t*u_tc');                        %
Linear position calculation
m = pwAddODE(m, 'x_tc', 'u_tc*cos(psi_t)-v_tc*sin(psi_t)');          %
Nonlinear position calculation
m = pwAddODE(m, 'y_tc', 'v_tc*cos(psi_t)+u_tc*sin(psi_t)');          %
Nonlinear position calculation
m = pwAddODE(m, 'psi_t', 'gam_t');
m = pwAddODE(m, 'psi_i', 'gam_i');

%% Observables
% m = pwAddY(m, rhs, ID, scalingParameter, errorModel)

m = pwAddY(m, 'gam_t', [], 'scale_gam_t_obs', '0.01*max(y)');
```

```
m = pwAddS(m, 'scale_gam_t_obs', 1, 'fix');
m = pwAddY(m, 'gam_i', [], 'scale_gam_i_obs', '0.01*max(y)');
m = pwAddS(m, 'scale_gam_i_obs', 1, 'fix');
```

ACKNOWLEDGEMENTS

I would like to thank my Lord and Savior Jesus Christ for my eternal salvation and for His unending grace during the various challenges of life during graduate study. I am thankful that this has been a time of significant growth not just in research and academics but also in my understanding of His mighty power and purpose for our lives.

I would like to express significant gratitude to my major professor, Dr. Brian Steward, for his outstanding guidance, wisdom, encouragement, and patience and for all of the learning opportunities that I have been given in my degree program. He has been an excellent role model to have at this stage of my life and career. I would also like to thank my other Program of Study committee members, Dr. Stuart Birrell and Dr. James Bernard, for the material that they taught in their engineering courses (tractor power and vehicle dynamics, respectively) and for the various contributions that they have made to my development both in and out of the classroom setting.

I also greatly appreciate Dr. Manoj Karkee for the guidance that he provided in the first year of my degree program as well as for the previous virtual reality and tractor-implement modeling and research work he performed that provided a strong foundation for me to build upon. I would also like to thank Zachary Kemp II of John Deere for the opportunity to collaborate on a very interesting and educational research assistantship and for all of the helpful knowledge and advice that he shared with me in that project.

I would like to thank the developers of PottersWheel at the University of Freiburg in Germany for creating and sharing a very useful toolbox for mathematical modeling and parameter identification tasks and for designing the toolbox general enough to be useful for non-biological system applications as well. I would especially like to thank Andreas Raue

for implementing the profile likelihood algorithm into the toolbox and for providing helpful feedback regarding my particular application.

Finally, I would like to thank my family for their love and encouragement and for the upbringing that created and continues to fuel my passion for engineering and agriculture.