Utilizing ANNs to Improve the Forecast for Tire Demand

A thesis presented to
the faculty of
the Russ College of Engineering and Technology of Ohio University

In partial fulfillment
of the requirements for the degree
Master of Science

Brent S. Taylor
May 2015
© 2015 Brent S. Taylor. All Rights Reserved.
This thesis titled
Utilizing ANNs to Improve the Forecast for Tire Demand

by

BRENT S. TAYLOR

has been approved for
the Department of Industrial and Systems Engineering
and the Russ College of Engineering and Technology by

Gary R. Weckman
Associate Professor of Industrial and Systems Engineering

Dennis Irwin
Dean, Russ College of Engineering and Technology
ABSTRACT

TAYLOR, BRENT S., M.S., May 2015, Industrial and Systems Engineering

Utilizing ANNs to Improve the Forecast for Tire Demand

Director of Thesis: Gary R. Weckman

This study is an initial attempt to investigate the relationship between economic factors and monthly tire sales, using artificial neural networks (ANNs) and comparing the results to stepwise regression. Data for this research were collected through a privately held tire warehouse located in Wheeling, West Virginia. Research has shown that artificial neural network models have been successfully applied to many real world forecasting applications. However, up to this date no research has been found using artificial neural networks and economic factors to predict tire demand. The first part of this research describes why the chosen economic factors were selected for this study and explains the initial methodology with results. The next stage of the research gives details on why the methodology was revised and also clarifies why Google Trends and additional mathematical inputs were applied to the study. The final research focused on separating the master database into three different categories based on selling percentages. The results of the study show that the artificial neural network models were capable of forecasting the number of high selling tires, with a validation technique, but were unable to be applied sufficiently for the medium and low selling products.
ACKNOWLEDGMENTS

I would first like to recognize Dr. Gary R. Weckman for sticking with me through
this entire process. It was a challenge, for me, to have a full time job and also focus on
finishing up my degree. However, Gary was extremely flexible through every step of the
way and I thank him. I’m not sure I would have been able to complete this journey
without his presence I would also like to thank Jan Weckman for allowing Gary to be
flexible and fun. They are both amazing, giving people and I cannot stress how much that
has helped along the way. Next I would like to thank my amazing wife Tiffany Taylor for
being patient, understanding, and supporting me to succeed. I would also like to
recognize Dr. William A. Young for his positive feedback and ideas along the way. He
was never was too busy to help, although he had a million other projects going on his life
at the time. Finally I would like to thank my parents, Brent and Sally Taylor, for raising
me to be a hard working gentleman. They always believed in me even when I did not
believe in myself. I thank them for their love and support through this entire process and
pushing me to be a better person, both personally and professionally.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>3</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>4</td>
</tr>
<tr>
<td>List of Tables</td>
<td>7</td>
</tr>
<tr>
<td>List of Figures</td>
<td>8</td>
</tr>
<tr>
<td>Chapter 1: Introduction</td>
<td>10</td>
</tr>
<tr>
<td>Tire Industry Background</td>
<td>10</td>
</tr>
<tr>
<td>Significance of Improved Accuracy</td>
<td>10</td>
</tr>
<tr>
<td>Main Tire Components</td>
<td>11</td>
</tr>
<tr>
<td>Problem with Current Research</td>
<td>12</td>
</tr>
<tr>
<td>ANNs</td>
<td>12</td>
</tr>
<tr>
<td>Explanation of Available Data</td>
<td>13</td>
</tr>
<tr>
<td>Tire Types and Differences</td>
<td>14</td>
</tr>
<tr>
<td>Bias Vs Radial Construction</td>
<td>15</td>
</tr>
<tr>
<td>Purpose/Objectives</td>
<td>16</td>
</tr>
<tr>
<td>Chapter 2: Literature Review</td>
<td>17</td>
</tr>
<tr>
<td>Current Research on Forecasting Tire Demand</td>
<td>17</td>
</tr>
<tr>
<td>Forecasting in Retail Markets</td>
<td>18</td>
</tr>
<tr>
<td>Demand Planning/Forecasting</td>
<td>19</td>
</tr>
<tr>
<td>Demand Forecasting Methods</td>
<td>19</td>
</tr>
<tr>
<td>Stepwise Regression</td>
<td>20</td>
</tr>
<tr>
<td>Artificial Neural Networks (ANNs)</td>
<td>21</td>
</tr>
<tr>
<td>Tire Demand and Cost Drivers</td>
<td>22</td>
</tr>
<tr>
<td>Model Performance Measures</td>
<td>23</td>
</tr>
<tr>
<td>Literature Review Summary</td>
<td>27</td>
</tr>
<tr>
<td>Chapter 3: Methodology</td>
<td>28</td>
</tr>
<tr>
<td>Data Collection and Compilation</td>
<td>30</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1. Model Variable Adjustments .......................................................... 36
Table 2. Additional Inputs - Google Trend Terms ........................................... 40
Table 3. Top Seller Metrics ........................................................................ 46
Table 4. Contributing Inputs ANN Compared to Stepwise Regression .......... 48
Table 5. Mid Seller Metrics ........................................................................ 52
Table 6. Low Seller Metrics ...................................................................... 54
Table 7. RSQ Metrics for all Three Selling Categories ................................. 56
Table 8. MSE Metrics for all Three Selling Categories ................................. 57
LIST OF FIGURES

Figure 1. Current data from National Tire and Wheel, Inc. ............................................. 14
Figure 2. Radial tire construction. .................................................................................... 16
Figure 3. Three layers of an Artificial Neural Network. .................................................. 22
Figure 4. Example of high error and good fit. ............................................................... 25
Figure 5. Example of low error and poor fit. ................................................................. 25
Figure 6. Example of high error and poor fit. ............................................................... 26
Figure 7. Example of low error and good fit ................................................................. 26
Figure 8. Methodology flow chart. ................................................................................ 30
Figure 9. Raw data report from the Mail Order Management (MOM) System. .............. 31
Figure 10. Example of part number sales per week. ...................................................... 32
Figure 11. Example of part number sales per half month. .......................................... 33
Figure 12. Example of part number sales per month.................................................... 33
Figure 13. Slope calculation for weekly data. ............................................................... 34
Figure 14. RSQ values for the initial 25 ANN models................................................... 37
Figure 15. Performance metrics for cumulative vs. monthly sales. ............................ 38
Figure 16. Comparison graph between cumulative prediction and total sold per month. 38
Figure 17. Cumulative estimate vs. actual sold............................................................. 39
Figure 18. Example of adding historical inputs............................................................ 41
Figure 19. Example of how moving average was calculated ........................................ 42
Figure 20. Final master database. .................................................................................. 43
Figure 21. Initial sensitivity analysis with 124 inputs. .................................................... 47
Figure 22. Final sensitivity with 12 inputs - top contributing inputs. ......................... 47

Figure 23. Linear regression model of estimate vs. actual for top sellers. .................. 49

Figure 24. Example of sales fluctuation for top sellers - five year history. ............... 50

Figure 25. Comparison of actual sales to ANN and stepwise estimates - top sellers....... 50

Figure 26. Linear regression model--estimate vs. actual for mid sellers...................... 51

Figure 27. Example of sales fluctuation for mid sellers - five year history. ............... 52

Figure 28. Comparison of actual sales to ANN and stepwise estimates - mid sellers. .... 53

Figure 29. Linear regression model--estimate vs. actual for low sellers...................... 53

Figure 30. Example of sales fluctuation for low sellers - five year history.................. 55

Figure 31. Comparison of actual sales to ANN and stepwise estimates - low sellers..... 55
CHAPTER 1: INTRODUCTION

Tire Industry Background

Consumer demand for rubber tires has been around ever since the late 1800s when Charles Goodyear developed a process to make rubber strong enough for commercialization. John Boyd Dunlap created the first pneumatic tire in the late 1880s (Bowlings, 2003). Since then, suppliers have been attempting to predict tire consumer purchasing trends. During the recent economic decline, the tire industry also saw a decrease in sales between the years of 2008 and 2009 (Global Industry Analysts, 2010). As a result of these inconsistent economic trends, suppliers of the tire industry have been researching ways to forecast tire demand. Improved forecasts allow suppliers to improve efficiency by being able to predict how to more accurately allocate their resources.

There are several variables involved that affect consumer tire consumption. Some theorized economic factors that affect demand are believed to be crude oil prices, gas prices, gross domestic product (GDP), interest rates, cost of steel, and the unemployment rate (Global Industry Analysts, 2010). Other factors the experts consider are average number of miles driven, the size of the tires, and previous years’ demand along with previous costs. By investigating all of these factors, this proposed study seeks to create a model to forecast tire sales.

Significance of Improved Accuracy

Having a more accurate forecast allows one to more efficiently allocate resources. Inventory can be both beneficial and costly, depending upon various circumstances. Companies carrying high volumes of stock tie up business capital, which decreases cash
flow and profitability (Oviamathi, 2014). The cost from carrying excessive inventory comes from several factors: the cost of the material itself; taxes being charged for holding the inventory; insurance to cover the inventory; and a payment for the building and overhead for the space needed to hold the inventory. In addition, most companies require labor to stock and pick the inventory (Anderson, 2004). However, although carrying inventory is expensive, not having the goods on hand causes customers to look elsewhere for the specific product they are trying to acquire. In the short term, this causes the loss of a sale; in the long term it could cost the company a customer (Schreibteder, 2010).

Without having the correct amount of inventory, the supplier loses money; therefore, if the forecast was improved this would benefit both the customer and the supplier. The supplier would be able to more efficiently allocate resources, and the customer would be able to get the product in a timely fashion.

**Main Tire Components**

A tire’s main component is rubber polymer; however, there is a vast array of other components involved to constantly improve their construction, and the ingredients continually increase with new knowledge and technology. Natural rubber from rubber trees and synthetic rubber (created using polymers) account for the majority of the tire. Tires also contain metal, fillers, softeners, various anti-aging ingredients, and curatives (Lindenmuth, 2006). Metal is used to create beads that resist blowing off, slipping, or breaking during the mounting process and during the tire’s lifetime on the rim. Metal is also used in radial tires to create cords, which strengthen and extend the life and reliability of the tire. The fillers within the tire are used to strengthen the rubber so that it
becomes more durable and reliable. The softeners are needed to allow the rubber to become more pliable and tacky so it can be shaped and formed more easily. The anti-aging chemicals allow the tire to resist degradation over time: If these were left out, the tire would begin to deteriorate at a much faster rate and could lead to a potentially dangerous situation on the highway. The curatives are used to accelerate the tire drying process and help all the components form together to make a strong rubber and a tough tire. (Lindenmuth, 2006).

**Problem with Current Research**

Currently, there has not been a specific forecasting method that incorporates the theorized factors that drive demand. Traditional models exist, but tend to be “one size fits all.” Most traditional models are linear and only rely on the previous term’s data to forecast (Mitrea, Lee, & Wu, 2009). Real world data can seldom be modeled accurately using basic linear regression. Therefore, some gaps exist that a non-linear model should be able to fill. Also, other forecasting methods do not take into consideration outside factors mentioned above such as crude oil prices, gas prices, unemployment rate, gross domestic product (GDP), interest rates, cost of steel, average number of miles driven, or the size of the tires. By using these researched and theorized factors, the results of the proposed study should be able to more accurately predict tire demand.

**ANNs**

Artificial neural networks (ANNs) have the ability to learn from numerical data in a method similar to a human brain. ANNs are able to learn and predict non-linear data better than or equal to traditional linear and non-linear regression techniques (Young II &
Weckman, 2010). ANNs have been used to predict everything from stock market fluctuations to football draft picks. ANNs have also been shown to be remarkable classifiers and pattern recognizers, according to Zhang (Zhang, Patuwo, & Hu, 1998). They are a great modeling tool that have unforeseen potential within any field of study, particularly in the field of engineering where mathematical modeling is highly useful. Since the neural network is a mathematical model, the data entered into the model has to be transformed into numerical values in order for the ANN to learn. Artificial neural networks do have one main downside; they lack transparency (Halgamuge, Poechmueller, & Glesner, 1995). Therefore, this tends to shy away users from the many applications neural networks can be designed for, and modelers seek out techniques with higher transparency such as linear regression. Although the lack of transparency is an issue to some researchers, the ability to achieve improved accuracy when forecasting is of higher importance to this study.

**Explanation of Available Data**

The data used in this case study comes from a tire warehouse that holds an average inventory of approximately 100,000 tires. The warehouse is a privately held business located in the panhandle of West Virginia. This company mainly focuses on light truck and Off Highway Vehicle (OHV) markets. The inventory and sales data has been collected for over ten years. However, the research being performed will focus on data from the past five years. The data shows part numbers, tire width, tire length, rim diameter, load range, biased ply or radial construction type, current amount in inventory, unit cost, received date, and sales date. This data is considered to be “lumpy” since there
are time periods where no demand or zero sales occur (Gutierrez, Solisb, & Mukhopadhyay, 2008). Therefore, some non-traditional methods, such as Artificial Neural Networks, should be better suited to predict this data. A screenshot of the data is shown in Figure 1. The aforementioned economic factors will be added to this data before analysis within the ANN.

<table>
<thead>
<tr>
<th>Part Number</th>
<th>Trans_date</th>
<th>QTY</th>
<th>Unit_Cvd</th>
<th>Total_Amt</th>
<th>resid</th>
<th>bin_Id</th>
<th>posdate</th>
<th>Name Brand</th>
<th>Height</th>
<th>Width</th>
<th>Rim Diameter</th>
<th>Load Radial or Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>07/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
<tr>
<td>4</td>
<td>07/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
<tr>
<td>6</td>
<td>08/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
<tr>
<td>8</td>
<td>08/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
<tr>
<td>10</td>
<td>08/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
<tr>
<td>12</td>
<td>08/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
<tr>
<td>14</td>
<td>08/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
<tr>
<td>16</td>
<td>08/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
<tr>
<td>18</td>
<td>08/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
<tr>
<td>20</td>
<td>08/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
<tr>
<td>22</td>
<td>08/01/11</td>
<td>-2</td>
<td>543.78</td>
<td>-2774.055</td>
<td>188</td>
<td>25875</td>
<td>11/02/11</td>
<td>BUGGER SUPER</td>
<td>44</td>
<td>19.5</td>
<td>15 C</td>
<td>Bias</td>
</tr>
</tbody>
</table>

Figure 1. Current data from National Tire and Wheel, Inc.

**Tire Types and Differences**

The tires of interest consist of light truck and automobile tires used by everyday drivers and off-road enthusiasts. Most tires in this study represent tires suited for light truck and Off Highway Vehicle (OHV) applications, which is a niche market in the tire industry. The tires of study range in height from 27-54 inches and vary in width from 9-22 inches. Every tire that will be a part of the testing and analysis is U.S. Department of Transportation (DOT) approved; therefore, they are street-legal. Non-street legal tires will not be considered in this study.
Bias Vs Radial Construction

Two types of street tires will be analyzed, bias ply and radial ply. These two types have very different characteristics, which could potentially be one of the main differences in determining demand. Compared to radial ply tires, bias ply tires are less reliable, since the sidewall and the tire all move as one piece (Lindenmuth, 2006). This creates poor wear characteristics since the tread of the tire is constantly changing during contact with the road. The bias tire also creates a rougher ride since every road irregularity is transferred to the automobile and the driver (Lindenmuth, 2006). Bias ply tires are beginning to become extinct in the automobile industry. However, a few tire companies still produce bias ply tires for automobiles due to their lower manufacturing costs. Radial ply tires, as shown in Figure 2, are more common today compared to 50 years ago, thanks to their increased durability and comfort (Discount Tire, 2012). Radial tires have a construction that allows the sidewall to flex or move separate from the tread. Therefore, they increase efficiency by creating a longer tread life and an improved fuel economy. The downside to radial tires is their high production and manufacturing cost, due to their complex nature.
Purpose/Objectives

The purpose of this study is to utilize ANNs to create a tire demand forecasting model using new theorized input variables. The theorized, size, historical, and economic variables will be integrated into one database and several ANNs will be created and tested. The best network developed will be compared to baseline techniques such as the Stepwise linear regression model.
CHAPTER 2: LITERATURE REVIEW

In order to provide background in the field of study, several topics will be reviewed. The process of determining demand drivers for tires can be quite complex; therefore, synthesis of the literature will be presented in order to verify and determine input variables. Artificial neural networks (ANNs) are also highly complex math models, so it is essential to determine how they work, along with how they have previously been used. Traditional forecasting techniques will also be examined, since they will be compared to ANNs later in the study.

Current Research on Forecasting Tire Demand

Current publicized research regarding tire sales forecasting is extremely limited. Experts in the tire industry predominantly rely on sales history and tribal knowledge when it comes to predicting the amount of tires to keep in inventory. Goodyear--a global leader in the manufacturing of tires and rubber products--relies on three different forecasting methods, each with separate metrics. Similar to the research in this study, Goodyear used historical data to predict future consumer demand through statistics, and found results leading to a greater than 50% inventory reduction (Miller & Liem, 2004). Some other studies have been conducted that look into tire buying behavior. Price, brand, performance, resistance, and tread depth are the top five factors that have been proven, in India, to influence tire consumer behavior (Natarajan, Soundararajan, & Jayakrishnan, 2013). Although this research will be conducted in the United States, these factors should be taken into consideration when attempting to forecast purchasing behavior. Forecasting tires has been shown to be difficult, due the fact that human psychology plays a crucial
role when it comes to how and when consumers decide to buy tires (Natarajan, Soundararajan, & Jayakrishnan, 2013). Another challenge in predicting tire sales lies in the tire salesman. According to Karelse, “80% of tire purchasing decisions are made by the salesman, not the consumer” (Karelse, 2013). This presents a new problem-- the possible parameters could be endless, based on promotional incentives, perceived quality, personal preference, word of mouth performance, etc... With that being said, forecasting is still possible, but a “one size fits all model” will be hard to achieve. Each situation will require a unique equation varying from a statistical method to a tribal knowledge approach (Karelse, 2013).

**Forecasting in Retail Markets**

Some research has already been completed regarding forecasting retail sales. The tire industry is also a retail market; therefore, some of the techniques used in predicting fashion retail markets may also be applied to the tire sales market. Retail markets rely heavily on traditional statistical methods, ranging from linear regression to ARIMA. These techniques are fast and easy to use, but they fall short when compared to Artificial Intelligence models (Na Liu, 2013). ANNs have been tested in retail markets and have returned higher accuracy metrics compared to traditional statistical methods in fashion retail and supermarkets (C. Frank, 2003). Another key metric found to influence any type of retail market comes from advertising and promotions. According to Bower, a key contributor to forecasting depends on seasonal sales (Bower, 2012). Therefore understanding seasonal marketing methods (holiday sales, yearly promotions or clearance
events, etc.) could be a potential factor that influences sales. However, this could also be a trend--historical data could be applied to predict sales if the promotions are seasonal.

**Demand Planning/Forecasting**

Demand planning is an essential tool used in industry to predict customer trends. Demand planning is about estimating the quantity, time, and location of future demand (Haberleitner, Meyr, & Taudes, 2010). When demand can be accurately predicted, it leads to customer satisfaction, along with allowing the service provider to avoid losing sales. Having the correct amount of inventory is essential, since inventory carrying cost can be 25%-55% of the inventory value on hand, according to Richardson (Richardson, 1995). Therefore, if demand can be accurately predicted, inventory levels can be decreased, since buffer stock inventory is needed only to cover inaccurate forecasts (Moon, Mentzer, Smith, & Garver, 1998). The ability to plan for future demand allows a company to reduce costs associated with purchasing inventory, such as ordering and logistics expenditures (Moon, Mentzer, & Thomas, 2000). Overall, the ability to plan for future demand provides numerous benefits to both the business and the customer.

**Demand Forecasting Methods**

Demand forecasting has the ability to save companies money by reducing inventory cost and increasing productivity (Hua & Zhang, 2006) (Gutierreza, Solisb, & Mukhopadhyay, 2008). Therefore, this section reviews some of the current research methods applied. Forecasting is either qualitative--which is a subjective approach by the forecaster--or quantitative, which is based on a calculated method (Reid & Sanders, 2010). Qualitative approaches are usually performed through surveys or by experts who
have a firm understanding of the data being forecasted. The expert approach is fast and flexible; however, the forecast is driven by one person or one group and lacks data-driven knowledge. The survey method is suitable for understanding preferences, but like most surveys it can be difficult to get customers to complete it, and developing a good questionnaire can be problematic (Reid & Sanders, 2010). Quantitative methods use historical data and are based on some form of statistical modeling. Time series demand planning can usually be broken into two categories, quantitative statistical methods and artificial intelligence methods (Wagner, Michalewicz, Schellenberg, Chiriac, & Mohais, 2011). Demands planning quantitative models usually consist of regression, time series (e.g., exponential smoothing, average, ARIMA), or a combination of methods (Hua & Zhang, 2006). Time series forecasting techniques are common in many distribution and manufacturing facilities. Linear regression is one of the most common and easiest methods used for predicting demand. However, since most real world data is nonlinear, traditional forecasting methods tend to produce large margins of error, and require some other forecasting technique to improve accuracy (Ghobbar & Friend, 2003). Adamoski et al, found ANNs to have an eleven percent increase in goodness of fit when compared to multiple linear regression models predicting water demand (Karapataki, 2010). This shows a classic example in which modeling with artificial neural networks can improve a forecast using real world data.

**Stepwise Regression**

Stepwise regression is a tool used to select the most significant inputs when several parameters are being analyzed in regression. Parameters are added and removed
one at a time and the model selects the top few based off the alpha value chosen. Details
can be controlled by adjusting the alpha value to the desired level. By performing a
Stepwise regression, the variables that show little significance to the model can be
removed allowing for a more concise and less complex model. This allows for easier
updating, since only the significant parameters need to be collected. Overall, Stepwise
regression is helpful when a large number of inputs are given to a model and the user
wants to see the significance of the most important ones. However, Stepwise regression
may not always return the highest R squared value when given several parameters,
compared to other data analysis tools (Barry, 2012).

**Artificial Neural Networks (ANNs)**

ANNs have proven themselves to be very useful, not only in academia but also in
industry. ANNs are non-linear modeling tools which make them beneficial at modeling
real world data (Kanungo, Arora, Sarkar, & Gupta, 2006) (Gaafar & Choueiki, 2000)
(Marroquinn, Cervantes, Flores, & Cabrera-Rios, 2009). However, determining how the
neural network obtains the results provided is problematic, since ANNs lack transparency
compared to traditional modeling techniques such as linear regression (Young II &
ANNs have three main layers: input, hidden, and one output layer, as shown in Figure 3
(Kanungo, Arora, Sarkar, & Gupta, 2006). The input layer consists of the attributes used
to estimate the final output. These are determined by the user. However, too many inputs
can create “noise” if the data has no significance to the model. Therefore, the user has to
make sure that the data entered is relevant. The amount of hidden layers is also
determined by the user. The hidden layers process the data, and the right amount is usually determined by trial and error analysis (Kanungo, Arora, Sarkar, & Gupta, 2006). The output layer processes data as well; however, it also displays the output, or result, of the artificial neural network. ANNs have been proven, mathematically, to be able to predict a wide array of analytical functions using MSE and $R^2$ values as a measure of calculation error (Marroquín, Cervantes, Flores, & Cabrera-Ríos, 2009). Although ANNs lack transparency they have the ability to be a more beneficial forecasting tool, compared to traditional statistical techniques (Mitrea, Lee, & Wu, 2009).

**Figure 3.** Three layers of an Artificial Neural Network.

**Tire Demand and Cost Drivers**

With new technology involved in the manufacturing of tires, there is much more than just rubber that affects the material costs. Cost drivers associated with tire demand
and consumer consumption have become quite challenging to predict. According to Global Industry Analysts (2010), tire demand is based on these economic cost drivers: “population growth, consumer confidence, discretionary spending, vehicle affordability, interest rates, credit availability, employment, and fuel prices”. Fisher and Vaidyanathan (2009) performed an assortment planning study using only three main tire factors: brand, size, and mileage warranty. However, they concluded that the research should have examined more outside factors than just the three they applied. The main ingredients needed to make a tire include filler, softeners, anti-degradant, curatives, and reinforcement materials such as cords and bead wires (U.S. DOT, National Highway Traffic Safety Administration, 2006). Most, if not all, of these materials should be considered when forecasting tire demand.

Model Performance Measures

To understand how well the model is performing, performance measuring indicators need to be used. Some basic statistical performance measures that can be applied are standard deviations, averages, and simple ratios. These are useful for determining the basics of whether or not the predictive model is correctly following the data. However, some better prediction and modeling metrics are found to better represent how good a model is at predicting trends and minimizing error.

One of the most common indicators for predicting trends is the r-squared ($R^2$ or RSQ) value, Equation 1 (Runkel, 2014). The RSQ value ranges from zero to one; the closer the value is to one, the more variation the model predicts or follows. The $R^2$ value
is able to capture the overall variance of the model, so its accuracy is relevant to the size of the model or dataset being used. Examples of models are shown below in the Figures.

\[
r^2 = 1 - \frac{SS \text{ Error}}{SS \text{ Total}} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}
\]

*Equation 1. RSQ formula.*

Mean Squared Error is a metric used to measure model performance. MSE is used to reflect the unbiased estimate of error variance. Shown in Equation 2, the MSE is calculated by dividing the error sum of squares by the degrees of freedom (Mielke jr., Berry, Landsea, & Gray, 1997). It is a highly popular metric used in variety of applications that are mainly focused around linear regression.

\[
\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2
\]

*Equation 2. MSE formula.*

The graphs below show the difference in the two performance metrics. The first graph, Figure 4 shows a High Error (MSE=64), but a Good Fit (RSQ=1).
Figure 4. Example of high error and good fit.

Figure 5 represents a Low Error (MSE=15), but a Poor Fit (RSQ=0.15).

Figure 5. Example of low error and poor fit

The High Error (MSE=82), Poor Fit (RSQ=0.09) is represented by Figure 6.
Lastly, a graphical representation showing “what good looks like” comes from Figure 7 with a Low Error (MSE=0.04) and Good Fit (RSQ=1). Ideally, this is the type of relationship that represents a good prediction model.
**Literature Review Summary**

The literature review demonstrates that very little work has been published in the field of forecasting tire demand. Basic linear mathematical models have weaknesses since most real world data is non-linear. Artificial neural networks have shown positive results in retail markets, and therefore the outcome of this research looks promising. The theorized demand drivers now have more validation, and the proposed research will work to provide evidence that these factors are either correct or incorrect at predicting tire sales.
CHAPTER 3: METHODOLOGY

This section covers the methods describing how utilizing ANNs to improve the forecast for tire demand will be derived. First, the data for all the inputs will need to be compiled and organized into the same time frame as the data from the tire warehouse. All the different databases for the theorized demand drivers—oil prices, GDP, interest rates, unemployment rates, etc.—will need to be found and integrated into one complete database. Each iteration of the data set will be tested by the ANN in order to help identify the significant inputs in the model. A “master database” will be created by combining pieces of the data collected; this united database will be used for the majority of the testing. The data will need to be pre-processed, which turns all the symbolic data into numerical values, so that the neural network is able to read the data. Once this is performed and the master database is created, several variations of the artificial neural network will be trained and tested. Additional forecasting methods will also be used in order to compare the ANNs to the traditional methods.

Different variations which involve changing the setup of the ANN will be performed, including the type of learning algorithm and the number of hidden layers and nodes. Some examples of the different learning algorithms are the multi-layer perceptron (MLP), the Levenburg-Marquart, and the genetic algorithm. The type of learning algorithm, along with the number of epochs and hidden layers, will be varied and recorded until a satisfactory model is created. The models that produce the most accurate results with the lowest mean squared error (MSE) and the highest R-Squared Value (RSQ) will then be investigated further by using sensitivity about the mean analysis.
Sensitivity analysis is a test that places the highest emphasis on the most significant inputs fundamental to the model (Hunter, Kennedy, Henry, & Ferguson, 2000). By using sensitivity analysis along with expert opinion, the study will hopefully be able to determine the main contributing variable inputs. A traditional method of forecasting will also be created from the master database. The error will be recorded and compared to the ANN. The results will then be analyzed and measured against the findings from the artificial neural network. Hopefully, the significant inputs will be consistent for all the models created. Figure 8, completes this section by showing the flow of the proposed methodology for both the artificial neural network and the traditional approach, using the new inputs.
Data Collection and Compilation

Data for the input parameters was collected from several sources and compiled to form a master database. A screen shot of the database is provided in Figure 9. Only data from 2007-2012 was acquired for this study, since 2007 was the last year National Tire & Wheel had available data from the mail order management system. The database consisted of over 65,000 data points over this five-year period. The raw data came from NTW’s mail order management system, also known as MOM.
Every sales transaction was recorded by National Tire & Wheel. This report was far from being ready for use—therefore; many iterations and data preprocessing techniques were required. Much of the time was spent on data extraction, since several parameters could be found in one column of data. The team originally decided to focus on two specific brands of tires. However, after several neural network models were created, the team decided to focus on one specific brand, Mickey Thompson (MT). MT Tires were the most popular tire brand for NTW at the time; therefore, Mickey was picked to be the sole focus of this study. NTW’s transaction report provided part numbers, tire size, demand, inventory, unit cost, order quantities, and load range of the tires that were all used as inputs in the master database. Number of Tires Sold also came from this database, and is the output this study is focusing on to predict. Now that the data has been preprocessed into a usable spreadsheet, some additional factors will need to be incorporated into the database.
Breaking Down the Data Weekly, Half Monthly, and Monthly

The data was separated out into three different segments throughout the research. The initial database was created using weekly data. Each unique part number was summed weekly (Example, shown in Figure 10), half-monthly (Figure 11), and finally, by month (Figure 12). A half-month in this study is defined as the total amount of tires sold before the fifteenth of the month and the total amount of tires sold after the fifteenth.

![Sales/Week](image)

*Figure 10. Example of part number sales per week.*
Making all the Data Weekly

Some of the data could only be found separated into months; therefore, a slope formula calculation was integrated into the database to separate this data into weekly
The study used the data from month to month to figure out the slope between the two. The increased or decreased slope between the two months was then used to predict the data for each week throughout the year, example shown in Figure 13. The formula designed to calculate weekly values for months with four weeks was:

*Calculation for Interest Rate: Previous interest rate + (¼ * slope between the two months)*

<table>
<thead>
<tr>
<th>Week</th>
<th>Month</th>
<th>Year</th>
<th>Monthly Unemployment Rate</th>
<th>Unemployment Rate (Formula)</th>
<th>Slope (Y)</th>
<th>Monthly Interest Rate</th>
<th>Interest rates (Formula)</th>
<th>Slope (Y)</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
<td>2007</td>
<td>4.50</td>
<td>4.50</td>
<td>0.20</td>
<td>3.75</td>
<td>3.75</td>
<td>0.20</td>
<td>0.5</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2007</td>
<td>4.58</td>
<td>4.58</td>
<td>0.75</td>
<td>3.75</td>
<td>3.75</td>
<td>0.75</td>
<td>1.25</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>2007</td>
<td>4.55</td>
<td>4.55</td>
<td>1</td>
<td>3.75</td>
<td>3.75</td>
<td>1</td>
<td>1.75</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>2007</td>
<td>4.53</td>
<td>4.53</td>
<td>1.25</td>
<td>3.75</td>
<td>3.75</td>
<td>1.25</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>2007</td>
<td>4.50</td>
<td>4.50</td>
<td>-0.04</td>
<td>4.72</td>
<td>4.72</td>
<td>-0.04</td>
<td>1.5</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>2007</td>
<td>4.47</td>
<td>4.47</td>
<td>1</td>
<td>4.72</td>
<td>4.72</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>2007</td>
<td>4.44</td>
<td>4.44</td>
<td>2</td>
<td>4.72</td>
<td>4.72</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>2007</td>
<td>4.42</td>
<td>4.42</td>
<td>2</td>
<td>4.72</td>
<td>4.72</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>2007</td>
<td>4.40</td>
<td>4.40</td>
<td>2</td>
<td>4.72</td>
<td>4.72</td>
<td>2</td>
<td>2.2</td>
</tr>
</tbody>
</table>

*Figure 13. Slope calculation for weekly data.*

The only change when a month contained five week was using 1/5 (calculation for one of five weeks), instead of using ¼ (calculation for one of four weeks). By using this formula, data was calculated for all 52 weeks of the year. This allowed every data set to be combined based on the week and year. A pivot table was created for all the data and each unique identifier was matched up to the corresponding date. This was done so that, when forecasting, the ANN would be able to more accurately pin point the forecast based on the week instead of the month or year.
**Economic Factors**

Seven economic factors were used as inputs to the model: crude oil price per barrel, gas price per gallon, cost of steel per hundredweight, interest rate, unemployment rate, average new car sales in the U.S., and average number of miles driven by consumers in the United States.

The crude oil and gas price data came from a database created by the U.S. Energy Information Administration (U.S. Energy Information Administration, 2012). The information provided from this Excel database consisted of daily price per barrel of crude oil, along with the price per gallon of regular gasoline, heating oil, kerosene, and propane from 1986 to the present. However, as mentioned earlier, the information used for this study only focused on price per barrel of crude oil and the price per gallon of regular gasoline, since those were thought to be some of the main contributors in order to predict tire demand. The annual car sales data came from the National Automobile Dealers Association (NADA) monthly reports. The economic factors, interest, and unemployment rates came from the United States Bureau of Labor and Statistics (Statistics, 2012). The cost of steel was applied from a database posted under the resources section on the Unarco Rack website (Unarco Rack, 2012). This was one of the most complete steel databases found for the study. It had weekly prices and data which dated back until January of 2004. The average miles driven estimate data came from the U.S. Department of Transportation (DOT) monthly reports (United States Department of Transportation, 2012). All of this data was pulled together into one master database so that all the parameters could be used as inputs within the artificial neural network. These factors
were all compiled and date-matched within the original database, which tried to predict weekly and monthly sales from National Tire & Wheel.

**Initial Trials and Models with Economic Factors Only**

The initial models from the proposed methodology did not show significant or helpful findings. These initial models were created using the seven economic factors mentioned, along with the tires sold, the date, and basic tire information. The basic tire information included the overall size, rim size, cost, and inventory level. The first couple attempts at this returned R squared values of around 7%. Several adjustments were made for the best R squared value that could be obtained from the original economic factors, with the resultant rate around 40%. Several models were tried and were adjusted, which are shown in Table 1 below. However, even with all these modifications the models still failed to provide helpful predictions.

Table 1.

*Model Variable Adjustments*

<table>
<thead>
<tr>
<th>Variables Changed While Creating</th>
<th>Variables Applied</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Models</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Learning Rules</strong></td>
<td>Levenberg-Marquardt, SVM, Probabilistic, MLP, GRNN</td>
</tr>
<tr>
<td><strong>Hidden Layers</strong></td>
<td>Ranged from 1-5</td>
</tr>
<tr>
<td><strong>Functions</strong></td>
<td>Linear, Sigmoid, Tanh</td>
</tr>
<tr>
<td><strong>Processing Elements</strong></td>
<td>Ranged from 1-50</td>
</tr>
<tr>
<td><strong>Forecast Frequency</strong></td>
<td>Weekly, Monthly, Half-Monthly</td>
</tr>
<tr>
<td><strong>Tire Brands</strong></td>
<td>Models created for 5 top selling brands</td>
</tr>
</tbody>
</table>
The study then modeled the data cumulatively. The results of predicting cumulatively returned an R squared value of 93%. All the initial model tests are shown in Figure 14 below. As shown, the cumulative reached the highest RSQ value, of the 25 models created in the study, thus far. However, predicting cumulative data led to a high mean squared error shown in Figure 15.

![Figure 14. RSQ values for the initial 25 ANN models.](image)

The graphs below are for the cumulative models for the top sellers. A comparison between five year cumulative models and monthly models is shown in Figure 16. Two different y-axis scales were applied to the graph to show the difference in the amount of tires the model attempted to predict. The large difference in the quantity of tires is one of the reasons for the higher MSE.
Figure 15. Performance metrics for cumulative vs. monthly sales.

<table>
<thead>
<tr>
<th>Performance</th>
<th>Cumulative</th>
<th>Total Sold per month</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>19064.93</td>
<td>128.97</td>
</tr>
<tr>
<td>RSQ</td>
<td>0.93</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Figure 16. Comparison graph between cumulative prediction and total sold per month.
The cumulative model had a much higher RSQ value, but also a much higher mean squared error, since the values that were trying to be predicted were also much higher. Glancing at the graph in Figure 17, would seem the research hit a home run. However, taking a closer look shows the model can be off by as many as 750 tires. A higher R squared for cumulative was expected. However, the ability to apply cumulative knowledge, as a decision-making tool, in order to predict the number of tires sold was deemed unhelpful for business application. Therefore, the researcher decided to take another step and collect new data from Google Trends, as well as adding in some additional mathematical inputs, which are mentioned below.

**Revised Methodology Additional Inputs-Google Trends**

The scope of the project was expanded and additional data from Google was inserted into the model as inputs. Google has a tool called Google Trends which looks
into the number of times a specific phrase or term is used (Google, Inc., 2013). The data from this tool is displayed weekly and dates back to 2004. The data is normalized and uses a 0-100 scale, with 100 being the most popular time a term was searched and 0 being the least popular time (Google, Inc., 2013). Twenty-two terms were searched and used as inputs in the ANN forecasting model. The terms chosen were based on the literature review research and off of expert opinions. Some of the specific terms looked into finding out when customers would be buying tires based off of specific Google search terms. The twenty-two terms listed in Error! Reference source not found. below were the ones used in the model.

Table 2.

**Additional Inputs - Google Trend Terms**

<table>
<thead>
<tr>
<th><strong>tire sales</strong></th>
<th><strong>diesel trucks for sale</strong></th>
<th><strong>beach</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>national tire and wheel</td>
<td>used diesel trucks</td>
<td>vacations</td>
</tr>
<tr>
<td>ntw</td>
<td>lifted trucks for sale</td>
<td>weather</td>
</tr>
<tr>
<td>mickey thompson tires</td>
<td>pro comp lift kits</td>
<td>snow</td>
</tr>
<tr>
<td>mt tires</td>
<td>rough country lift kits</td>
<td>ice</td>
</tr>
<tr>
<td>pro comp tires</td>
<td>hotels</td>
<td>rain</td>
</tr>
<tr>
<td>mud tires</td>
<td>tickets</td>
<td>the stock market</td>
</tr>
<tr>
<td>all terrain tires</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All of these terms were thought to possibly have some impact on the amount of tires being sold. By looking at how many times these terms were “Googled” the researchers were hoping to obtain information about the influence of weather conditions,
the amount of expendable money customers had by looking at vacations, and market fluctuations. The research also looked at whether the number of hits on search terms about tires or trucks had an influence on if the customers would be buying tires.

Mathematical Inputs

Three additional mathematical inputs were derived as inputs to the model. These included moving averages of all the economic factors, the cumulative sales of all the tires sold per month, and the delta—or difference—between the sales of tires each month. Inputs were also created from historical data looking back at either the prior month or the prior six months’ actual sales, shown in Figure 18. The research also added inputs for the previous six months’ average sales, the difference between the predicted prior month to the prior six months’ actual sales, and the difference between the prior month to six months’ average sales.

<table>
<thead>
<tr>
<th>Total Sold per Month</th>
<th>Prior Month</th>
<th>Prior 2 Month</th>
<th>Prior 3 Month</th>
<th>Prior 4 Month</th>
<th>Prior 5 Month</th>
<th>Prior 6 Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>12</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>12</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>8</td>
<td>17</td>
<td>12</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>20</td>
<td>8</td>
<td>17</td>
<td>12</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>32</td>
<td>20</td>
<td>8</td>
<td>17</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>24</td>
<td>25</td>
<td>32</td>
<td>20</td>
<td>8</td>
<td>17</td>
<td>12</td>
</tr>
</tbody>
</table>

*Figure 18. Example of adding historical inputs.*
Moving Average of Inputs

Along with all of the inputs the study also wanted to ascertain how significant of a contribution the moving average of all the inputs would perform in the model. Therefore, a three month and three-half month moving average was calculated for all of these trends, and was also used as an input to the dataset. Figure 19, shows how sales history was used to create moving averages. The two through six month moving average was collected for sales, in order to give the model sufficient sales history to learn from. Figure 19 shows how the four highlighted sales history equates to an average of 14.25 for that specific tire for the previous four months.

![Table showing data on sales per month and moving averages](image)

*Figure 19. Example of how moving average was calculated.*

In addition to the sales history data and the tire metrics, moving average calculations made for a total of 66 inputs from Google trends and 21 inputs from the economic factors. The master database is shown in Figure 20.
Revised Models One Database

After the additional Google trend and mathematical inputs were entered into the database, the study moved forward and began to build artificial neural networks to predict the number of tires sold. The initial results returned an R squared value of 40%. This was almost identical to the results before adding in all the additional data. Therefore, the researchers decided to break the data up into three separate databases, all focusing on one high-volume tire brand.

Separate Tire Databases with High Volume Tire Brand

National Tire and Wheel sells and stocks approximately 10-20 different tire brands. The research performed in this study focused on the top sellers. During the initial methodology, three master databases were created with all the aforementioned factors
aligned with each brand. Each database had close to 20,000 rows of data with ANN’s built and tested on each dataset. The database with the top results was the one chosen to perform a more detailed analysis and breakdown. National Tire stocks 136 different sizes of the particular brand focused on here, Mickey Thompson. When broken down symbolically within the neural network, this added another 136 inputs to the model, since every part number is unique. The research wanted to make sure the network was able to differentiate between each specific part number, hoping this unique identifier would play a key role in predicting sales trends.

**All Data**

Since running the model with all the data was achieving less than pleasing results, a decision was made to break the data up into categories. These three categories were based on the amount sold. Top sellers were the tires sold each month, for 100 percent of the time, for at least a year. The middle/average sellers were the ones sold each month for 50-75 percent of the timeline. The low model consists of tires that were sold less than 50 percent of the time.

**The Final Models**

Over 50 ANN models were created throughout this research. The best results in this research came from running three hidden layers and using a momentum learning rule. The number of hidden neurons applied were 20-15-10 (in that order) for the first, second, and third hidden layers. According to Gately, the number of neurons is not linked to a critically improving accuracy (Gately, 1996). The number chosen for this study were based off of several trial-and-error variations. The transfer function was a TanhAxon for
each hidden layer and the output used a LinearAxon transfer. These parameters were found to create the best results.

**Sensitivity Analysis**

Sensitivity analysis is a way to reduce noise, remove uncertainty and identify the most significant inputs within a predictor model. Running a sensitivity analysis helps to determine which parameters are helpful and which parameters can be removed (Wagner S., 2007). A sensitivity analysis is performed by assigning a rating to each of the predictor inputs. There are many theories as to how this should be performed (Hunter, Kennedy, Henry, & Ferguson, 2000). However, in this study the sensitivity about the mean test was the one built into the neural network software, using two standard deviations and fifty steps per side. Once the analysis was performed, insignificant parameters were removed. This usually leads to better metrics and an improved model that is less complex and cumbersome to populate. In this study, it was imperative to reduce the amount of inputs required. This would allow this tool to require less time and manpower to gather and populate. For the top seller model, performing sensitivity about the mean analysis reduced the number of inputs by 84%, leaving only the critical inputs as the final predictors.

**Top Sellers**

The definition of our top selling part numbers are tires that were sold from National Tire and Wheel at least once per month for an entire year. Thus, they would have a sales percentage for the year of 100%, or 1.
The initial step in predicting the top part numbers was to run a test with all 124 inputs from Figure 21. These preliminary results looked promising, as the R-Squared value was at 58%. The model was reduced three times after running sensitivity about the mean analysis and removing the inputs with the least significance. Eliminating inputs reduces noise within the model and usually results in a higher accuracy and better performance. The first reduction eliminated 86 inputs. The model was reduced until the performance of the model started declining. The first three gradually improved performance with the third returning an RSQ value of 66%. The fourth reduction resulted in an RSQ Value of 26%. Table 3, displays the metrics for a comparison between all the iterations for the artificial neural network models compared to the Stepwise and standard linear regression. As shown, the third iteration of the ANN returned the best R squared value. However, the Stepwise regression model returned a marginally lower mean squared error. Overall, the best model for predicting top selling part numbers turned out to be the third reduced model, with a total of fourteen parametric and attribute inputs, plus the six part number inputs shown in Figure 22 below.

Table 3.

Top Seller Metrics

<table>
<thead>
<tr>
<th>Top Sellers</th>
<th>ANN First Iteration</th>
<th>ANN Second Iteration</th>
<th>ANN Third Iteration</th>
<th>ANN Fourth Iteration</th>
<th>Stepwise Regression</th>
<th>Standard Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSQ</td>
<td>58%</td>
<td>60%</td>
<td>66%</td>
<td>26%</td>
<td>49%</td>
<td>48%</td>
</tr>
<tr>
<td>MSE</td>
<td>273.3</td>
<td>168.18</td>
<td>164.46</td>
<td>176.56</td>
<td>163.96</td>
<td>165.42</td>
</tr>
</tbody>
</table>
Figure 21. Initial sensitivity analysis with 124 inputs.

These inputs, from Figure 22, within the ANN were able to predict the top one percent of tire sales with an RSQ value of 66%, which turned out to be the best model for predicting the top selling part numbers.

Figure 22. Final sensitivity with 12 inputs - top contributing inputs.
Table 4 below shows the identified top contributing inputs from the Sensitivity Analysis performed with the artificial neural network compared to the top contributing inputs identified within the Stepwise regression Model. Both models used a few of the same inputs, but the significant weight or contribution did not weigh heavy from one single source. All the inputs provided valuable contributions to the models.

Table 4.

*Contributing Inputs ANN Compared to Stepwise Regression*

<table>
<thead>
<tr>
<th>ANN - Sensitivity Analysis</th>
<th>Stepwise With All Inputs</th>
<th>Stepwise With ANN-Sensitivity Analysis Inputs Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Miles Driven</td>
<td>Cuml3MMA</td>
<td>Average Miles Driven</td>
</tr>
<tr>
<td>Average Miles Driven 3HMMA</td>
<td>Cuml6MMA</td>
<td>Dates</td>
</tr>
<tr>
<td>Dates</td>
<td>delta1</td>
<td>mud tires3MMA</td>
</tr>
<tr>
<td>diesel trucks for sale3HMMA</td>
<td>Gas Price</td>
<td>Sales Percentage</td>
</tr>
<tr>
<td>Gas Price 3MMA</td>
<td>Gas Price 3MMA</td>
<td>tickets3MMA</td>
</tr>
<tr>
<td>Load Range(E)</td>
<td>InvTot Formula</td>
<td>Unemployment rate</td>
</tr>
<tr>
<td>mickey thompson tires</td>
<td>pro comp lift kits</td>
<td></td>
</tr>
<tr>
<td>mud tires3MMA</td>
<td>Sales Percentage</td>
<td></td>
</tr>
<tr>
<td>Prior Classification</td>
<td>Six</td>
<td></td>
</tr>
<tr>
<td>Sales Percentage</td>
<td>the stock market3HMMA</td>
<td></td>
</tr>
<tr>
<td>tickets3MMA</td>
<td>vacations3MMA</td>
<td></td>
</tr>
<tr>
<td>Unemployment rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Width</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The graphs below represent a specific top selling part number. All part numbers predicted can be found in the Appendix. These are intended to show the fluctuation in the
number of tires sold each month. These graphs also show the results from the ANN and the Stepwise regression. Figure 24, shows the trends over the five-year span. Figure 25 shows a close-up view of how the artificial neural network and Stepwise regression model predicted sales, compared to actual sales. Figure 23, was placed in this text to demonstrate how a basic linear regression model fit the data. The mean square in R squared values for these graphs are shown above in Table 3. The high variation and unpredictability from sales, even from the top selling part numbers, was a main reason fitting a model to this type of behavior was difficult.

*Figure 23. Linear regression model of estimate vs. actual for top sellers.*
Mid Sellers

With the top models bringing the best and most useful results of the research thus far, the research indicated to try the same theory on the mid/average sellers. The same model was utilized, with three hidden layers and a momentum learning rule. The initial run returned results with an RSQ value of 19%. The model was reduced to eliminate some noise, with expectations of the model improving. Once the sensitivity about the mean was run, 64 of the inputs creating noise were eliminated, leaving the more
influential inputs to train the reduced mid model. The reduced model returned similar results with a small improvement, an RSQ of 22%. A third reduction iteration was performed and returned RSQ results of 20%. Therefore, the second model was chosen when comparing this model to Stepwise regression, the latter actually outperformed the artificial neural network by a slim margin; a 4% better R squared value. One of the reasons this could have occurred is the algorithm. When training an artificial neural network model, a cross validation data set is required. However, in Stepwise regression, this cross validation data is not needed; therefore, this allows more data for training, which would lead to better results. Figure 26, shows estimate versus actual from the results of a standard linear regression.

All the results for the mid selling tires are shown in Table 5. Overall, the model showed promising results for the top sellers, but the results were rather lackluster in both, Stepwise regression and the ANN model, for the average, or mid-selling tires.
Table 5.

**Mid Seller Metrics**

<table>
<thead>
<tr>
<th>Mid Sellers</th>
<th>ANN First Iteration</th>
<th>ANN Second Iteration</th>
<th>ANN Third Iteration</th>
<th>Stepwise Regression</th>
<th>Standard Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSQ</td>
<td>19%</td>
<td>22%</td>
<td>20%</td>
<td>24%</td>
<td>23%</td>
</tr>
<tr>
<td>MSE</td>
<td>36.96</td>
<td>34.91</td>
<td>37.49</td>
<td>31.41</td>
<td>31.35</td>
</tr>
</tbody>
</table>

As described above for the top sellers a specific part number was also chosen for the mid sellers to represent the fluctuation in tire sales. The results for six months timeframe for both Stepwise and artificial neural network are shown in Figure 27. Figure 28 shows a more concise view of the results.

![Part Number(MTMTZ5268)](image)

*Figure 27: Example of sales fluctuation for mid sellers - five year history.*
Low Sellers

The least popular (or low selling tires) were the ones that only sold 25% of the time or less each month for one of the five years. The linear regression model shown in Figure 29, displays the estimate versus actual line. One thing to notice in this graph is how the low sellers mostly bunch around the range of zero and four. However, the amount sold per month has shown to jump above 40.

Figure 28: Comparison of actual sales to ANN and stepwise estimates - mid sellers.

Figure 29: Linear regression model--estimate vs. actual for low sellers.
The same model was applied again that brought the aforementioned good results from the top sellers. The model was retrained with the new data. The RSQ value for the low model started off at 19%. The model was reduced until the results started decreasing; the final low seller model that was picked had an R squared value of 21%. Similar findings came from the low sellers that were found with the mid sellers, where the R squared value for Stepwise regression was slightly better than the artificial neural network, Table 6. As stated earlier, this is thought to be influenced by the additional data coming from the cross validation set that could be used as testing data in Stepwise regression. As shown in Figure 30 below, the high variability in demand from consumers was the main challenge the researchers were faced with. Figure 31 shows a closer look of how the sales for the low-selling tires have a lumpy demand. Sales can be obsolete for months with occasional sales spikes in between. Both the artificial neural network and the Stepwise regression found the best fit when averaging out the amount of tires sold per month. More future research may lead to better understanding consumer behavior. However, within the scope of using the theorized economic factors and Google trends, no strong correlation was found between those parameters and tire sales.

Table 6.

Low Seller Metrics

<table>
<thead>
<tr>
<th>Low Sellers</th>
<th>ANN First Iteration</th>
<th>ANN Second Iteration</th>
<th>ANN Third Iteration</th>
<th>Stepwise Regression</th>
<th>Standard Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSQ</td>
<td>19%</td>
<td>21%</td>
<td>18%</td>
<td>25%</td>
<td>22%</td>
</tr>
<tr>
<td>MSE</td>
<td>11.09</td>
<td>10.55</td>
<td>10.73</td>
<td>11.08</td>
<td>11.24</td>
</tr>
</tbody>
</table>
Figure 30. Example of sales fluctuation for low sellers - five year history.

Figure 31. Comparison of actual sales to ANN and stepwise estimates - low sellers.
CHAPTER 4: FINAL RESULTS

The final results from this study for all selling categories are shown in Table 7 and Table 8. The data set returning the highest R squared value came from the third iteration of the high selling part numbers. The model performance for the mid and low selling tires was almost identical. Occasionally, the regression models would outperform the artificial neural network. However, as mentioned previously, this could be caused by the algorithm requiring cross validation whereas the regression models were able to utilize this data for analysis. Overall, for the top sellers, the artificial neural network was by far the best at predicting sales and most likely capable of being applied to the business.

Table 7.

\textit{RSQ Metrics for all Three Selling Categories}

<table>
<thead>
<tr>
<th>RSQ</th>
<th>Top Sellers</th>
<th>Mid Sellers</th>
<th>Low Sellers</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN First Iteration</td>
<td>58%</td>
<td>19%</td>
<td>19%</td>
</tr>
<tr>
<td>ANN Second Iteration</td>
<td>60%</td>
<td>22%</td>
<td>21%</td>
</tr>
<tr>
<td>ANN Third Iteration</td>
<td>66%</td>
<td>20%</td>
<td>18%</td>
</tr>
<tr>
<td>ANN Fourth Iteration</td>
<td>26%</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Stepwise regression</td>
<td>49%</td>
<td>24%</td>
<td>25%</td>
</tr>
<tr>
<td>Standard Regression</td>
<td>48%</td>
<td>23%</td>
<td>22%</td>
</tr>
</tbody>
</table>
Table 8.

### MSE Metrics for all Three Selling Categories

<table>
<thead>
<tr>
<th>MSE</th>
<th>Top Sellers</th>
<th>Mid Sellers</th>
<th>Low Sellers</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN First Iteration</td>
<td>273.3</td>
<td>36.96</td>
<td>11.09</td>
</tr>
<tr>
<td>ANN Second Iteration</td>
<td>168.18</td>
<td>34.91</td>
<td>10.55</td>
</tr>
<tr>
<td>ANN Third Iteration</td>
<td>164.46</td>
<td>37.49</td>
<td>10.73</td>
</tr>
<tr>
<td>ANN Fourth Iteration</td>
<td>176.56</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Stepwise regression</td>
<td>163.96</td>
<td>31.41</td>
<td>11.08</td>
</tr>
<tr>
<td>Standard Regression</td>
<td>165.42</td>
<td>31.35</td>
<td>11.24</td>
</tr>
</tbody>
</table>
CHAPTER 5: CONCLUSIONS

From this research, artificial neural networks, along with other predictor models, were shown to have limitations when predicting tire sales for this case study, at least when based on the defined economic factors that were used as inputs. For the top-selling tire category, the model could be used to predict sales and required inventory levels with the help of expert opinion validating the models predictions. The mid- and low-selling models could be used as a forecasting reference tool for the expert buyer. For the most part, the mid- and low-selling models tended to predict sales by using an average of sales history figures. Therefore, these models could be used as a guide to help businesses react to future demand, keeping in mind that some fluctuation is usually going to occur. Applying this research to nationwide tire sales could show better results by having an increased market size instead of focusing on a privately held business. However, with so much variability within the sales of this private business, it does make forecasting and predicting consumer behavior extremely difficult. According to Patrick Bower, a forecasting guru, “It would be great if there were a forecasting algorithm that reads consumers’ minds, but there isn’t.” (Bower, 2012) With that said, businesses have to make decisions on historical data, expert opinion, and relative parameters contributing to or relating to consumer buying behavior. Although the results in this research are not as strong as one would have hoped, great lessons were learned along the way, and some new sales prediction ideas were given for businesses in their future endeavors.
CHAPTER 6: FUTURE RESEARCH

The future of this research is almost limitless, as long as one wants to focus on continuous improvement with an endless scope. There are several directions the research could potentially take. First, more inputs and parameters could be added and used as predictors. A new category focused on the time a tire was introduced to the market could be added, for example--this is based on the assumption that newer tires tend to sell more. Another approach would be for the business to better understand large dealer buying patterns. Dealers are companies who buy from the researched warehouse in order to supply their customer or stock their own warehouse for future demand. Some of the spikes within the data could be the result of these dealers. Therefore, having a better understanding of this buying behavior could potentially categorize tires into different subgroups that could be used for prediction. Models could also be created by having the experts categorize tires by market trends based on what they’re hearing from consumers and marketing forecasters. Finally, these models could potentially work better when looking at nationwide tire sales as a whole. This data may be hard to obtain, but the larger market could be driven more by economic factors than privately held businesses. In conclusion, there are a number of logical options for future research and the boundaries are nearly limitless.
REFERENCES


APPENDIX: INDIVIDUAL PART NUMBER DEMAND PREDICTION

**Part Number (MTATZ5373)**

**Part Number (MTFCIICEP13274)**
Part Number (MTCLAWR5756)

Part Number (MTCLAWR5766)

Part Number (MTCLAWR5766)
Part Number (MTCLA\textit{W}R5768)

Part Number (MTCLA\textit{W}R5768)

Part Number (MTCLA\textit{W}R5776)
Part Number (MTMCCEP23253)

Part Number (MTMCCEP23264)

Part Number (MTMCCEP23264)
Part Number (MTSTZ50631)

Part Number (MTSTZ50631)

Part Number (MTSTZ50740)
Part Number (MTTTC5863)

- Total Sold per month
- ANN Pred
- Stepwise Pred

1 2 3 4 5 6