# TRUCK FORECASTING WITH TIME SERIES ANALYSIS: A CASE STUDY OF THE BLUE WATER BRIDGE

University of Wisconsin – Milwaukee
Paper No. 11-5

National Center for Freight & Infrastructure Research & Education
College of Engineering
Department of Civil and Environmental Engineering
University of Wisconsin, Madison

*Authors:* Jing Mao and Alan J. Horowitz
Center for Urban Transportation Studies
University of Wisconsin – Milwaukee

*Principal Investigator:* Alan J. Horowitz
Professor, Civil Engineering and Mechanics Department, University of Wisconsin – Milwaukee

December 22, 2011

# Truck Forecasting with Time Series Analysis: A Case Study of the Blue Water Bridge

## INTRODUCTION

This document contains images of all slides in a course module about the use of time series techniques for truck forecasting. The techniques are illustrated with data from the Blue Water Bridge between Michigan and Ontario. This presentation is available upon request to Alan Horowitz, horowitz@uwm.edu.

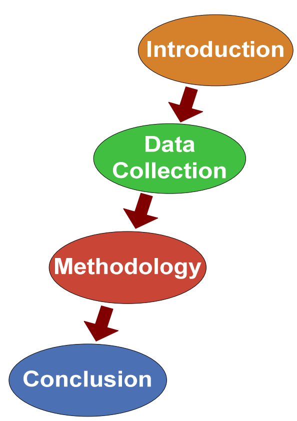
# Truck Forecasting with Time Series Analysis: A Case Study of the Blue Water Bridge

Prepared by
Jing Mao
Alan J. Horowitz



## Outline

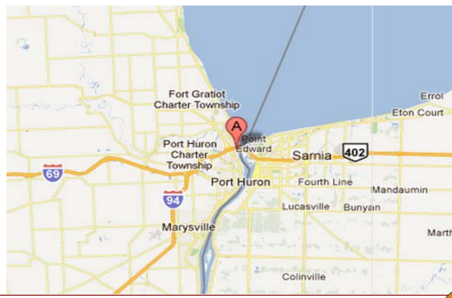
- Introduction
- Data Collection
- Methodology
- Conclusion



# Blue Water Bridge

## Location

The Blue Water Bridge spans the Saint Clair River, and carries international traffic between Port Huron, Michigan and Point Edward and Sarnia, Ontario. Located near interchange of I-94 and I-69, the bridge forms a critical gateway linking Canada and the United States.






# Blue Water Bridge

## Lane characteristics

The original Blue Water Bridge, opened in 1938 and renovated in 1999, is a three-lane westbound bridge. The second Blue Water Bridge, which carries three lanes of eastbound traffic, is an impressive modern bridge opened in 1997.

## Purpose of the Study

- Forecasting the eastbound and westbound monthly truck volume of blue water bridge from 2011 to 2013
- Applying the different time series models and select the best one for forecasting

Time Series Analysis

Introduction

## Data Source

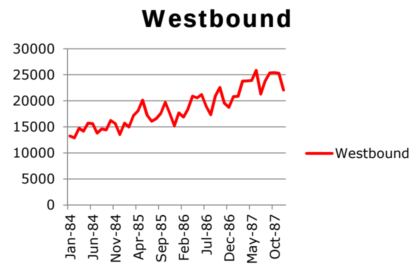
- Dependent variables
  Blue water bridge eastbound/ westbound truck volume
- Independent variables
  - Michigan population (why freight moves)
  - Ontario population (why freight moves)
  - U.S. GDP and population (why freight moves)
  - Approximate Michigan GDP (derived from U.S. GDP by the proportion of Michigan population and U.S. population)
  - U.S. all grades all formulations retail gasoline fuel price (major cost of freight)
  - North American Free Trade Agreement (NAFTA) (why freight moves)
  - September 11 attacks (why freight does not move)

Time Series Analysis

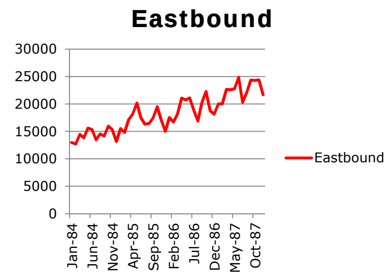
Data Collection

# Descriptive Statistic of Dependent Variables

- Westbound and eastbound truck volume data from Jan 1984 to Dec 2010 by each month




Westbound truck volume from Jan 1984 to Dec 1987



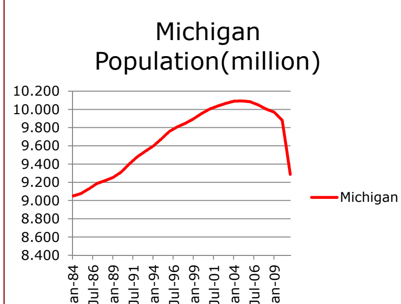

Eastbound truck volume from Jan 1984 to Dec 1987

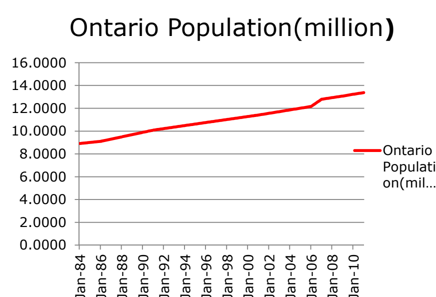
Time Series Analysis

Data Collection

# Descriptive Statistic of Independent Variables

- Michigan and Ontario population data from Jan 1984 to Dec 2010 by each month



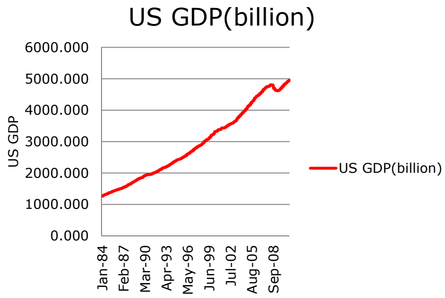



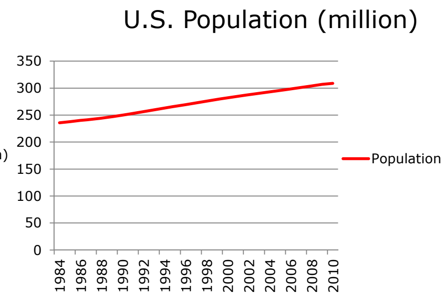

Time Series Analysis

Data Collection

# Descriptive Statistic of Independent Variables

- US GDP data from Jan 1984 to Dec 2010 by each month and yearly U.S. population data




Time Series Analysis

Data Collection

# Approximate Michigan GDP

- Computing the ratio of Michigan population and U.S. population
- Computing Michigan GDP by applying the ratio to Michigan GDP and U.S. GDP

| Year | Michigan Populati | U.S. Population | Ratio | U.S. GDP | Michigan GDP |
|---|---|---|---|---|---|
| 1984 | 9.049 | 235.825 | 0.038372 | 3,930.9 | 150.835 |
| 1985 | 9.076 | 237.924 | 0.038147 | 4,217.5 | 160.883 |
| 1986 | 9.128 | 240.133 | 0.038012 | 4,460.1 | 169.539 |
| 1987 | 9.187 | 242.289 | 0.037918 | 4,736.4 | 179.593 |
| 1988 | 9.218 | 244.499 | 0.037702 | 5,100.4 | 192.293 |
| 1989 | 9.253 | 246.819 | 0.037489 | 5,482.1 | 205.519 |
| 1990 | 9.311 | 249.623 | 0.0373 | 5,800.5 | 216.360 |
| 1991 | 9.4 | 252.981 | 0.037157 | 5,992.1 | 222.648 |
| 1992 | 9.479 | 256.514 | 0.036953 | 6,342.3 | 234.368 |
| 1993 | 9.54 | 259.919 | 0.03670 | | |

Michigan GDP = Ratio * U.S. GDP

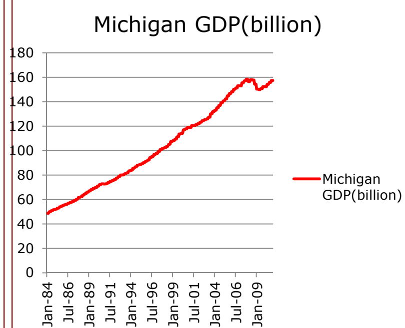
Ratio = Michigan population / U.S. population

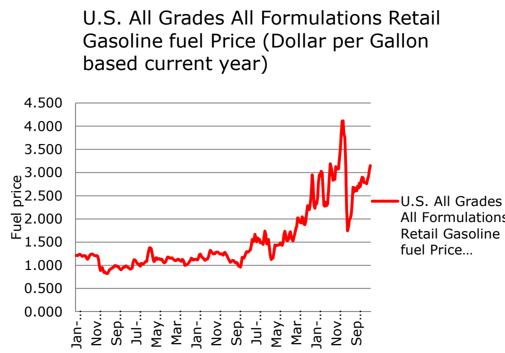
Time Series Analysis

Data Collection

# Descriptive Statistic of Independent Variables

- Michigan GDP data and U.S. all grades all formulations retail gasoline fuel price data from Jan 1984 to Dec 2010 by each month




Time Series Analysis

Data Collection

# Other Data

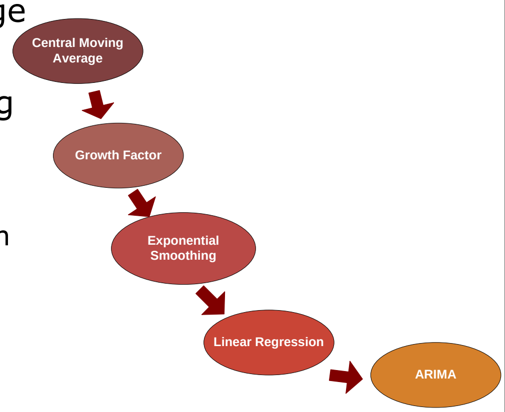
- NAFTA
  - The North American Free Trade Agreement or NAFTA is an agreement signed by the governments of Canada, Mexico, and the United States, creating a trilateral trade bloc in North America. The agreement came into force on January 1, 1994. It superseded the Canada – United States Free Trade Agreement between the U.S. and Canada.
- September 11
  - The September 11 attacks could also be a factor to influence the truck volume within that month.

Time Series Analysis

Data Collection

# Time Series Models

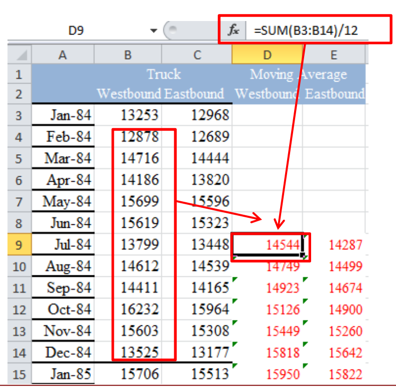
- Central Moving Average
- Growth Factor
- Exponential Smoothing
- Linear Regression
- ARIMA
  - o Box-Cox Transformation





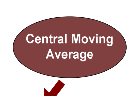

# Central Moving Average

- Central moving average is a moving average such that time period is at the center of the N time periods used to determine which values to average.



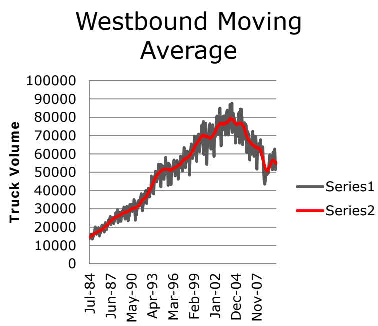
D9 =SUM(B3:B14)/12

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | | Truck | | Moving Average | |
| 2 | | Westbound | Eastbound | Westbound | Eastbound |
| 3 | Jan-84 | 13253 | 12968 | | |
| 4 | Feb-84 | 12878 | 12689 | | |
| 5 | Mar-84 | 14716 | 14444 | | |
| 6 | Apr-84 | 14186 | 13820 | | |
| 7 | May-84 | 15699 | 15596 | | |
| 8 | Jun-84 | 15619 | 15323 | | |
| 9 | Jul-84 | 13799 | 13448 | 14544 | 14287 |
| 10 | Aug-84 | 14612 | 14539 | 14749 | 14499 |
| 11 | Sep-84 | 14411 | 14165 | 14923 | 14674 |
| 12 | Oct-84 | 16232 | 15964 | 15126 | 14900 |
| 13 | Nov-84 | 15603 | 15308 | 15449 | 15260 |
| 14 | Dec-84 | 13525 | 13177 | 15818 | 15642 |
| 15 | Jan-85 | 15706 | 15513 | 15950 | 15822 |

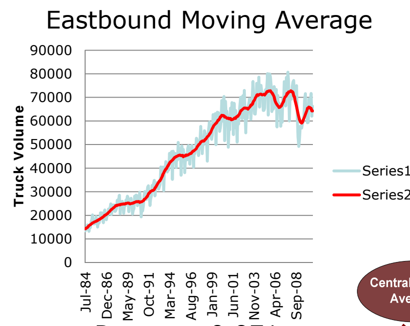


# Central Moving Average

- Westbound and Eastbound moving average data from Jul 1984 to Aug 2010




R square 0.956




R square 0.971

Central Moving Average

Methodology

Time Series Analysis

# Seasonal adjustment factor

- Seasonal adjustment factor can be used to improve accuracy of truck volume forecasts

| | Mean of Truck Volume | | Mean of Moving Average | | Seasonal Adjustment Factor | |
|---|---|---|---|---|---|---|
| **Month** | Westbound | Eastbound | Westbound | Eastbound | Westbound | Eastbound |
| Jan | 47598 | 43333 | 50923 | 47456 | 0.9347 | 0.9131 |
| Feb | 46702 | 43269 | 51043 | 47612 | 0.9150 | 0.9088 |
| Mar | 52219 | 48566 | 51174 | 47778 | 1.0204 | 1.0165 |
| Apr | 51219 | 47691 | 51307 | 47948 | 0.9983 | 0.9946 |
| May | 53322 | 49808 | 51434 | 48114 | 1.0367 | 1.0352 |
| Jun | 53390 | 50476 | 51560 | 48272 | 1.0355 | 1.0457 |
| Jul | 45114 | 42391 | 50097 | 46466 | 0.9005 | 0.9123 |
| Aug | 51872 | 49466 | 50221 | 46615 | 1.0329 | 1.0612 |
| Sep | 51958 | 49114 | 50344 | 46764 | 1.0320 | 1.0502 |
| Oct | 54683 | 51480 | 50493 | 46940 | 1.0830 | 1.0967 |
| Nov | 51412 | 48216 | 50633 | 47110 | 1.0154 | 1.0235 |
| Dec | 43788 | 41711 | 50772 | 47276 | 0.8624 | 0.8823 |

Central Moving Average

Methodology

Time Series Analysis

## Growth Factor

Linear growth:

$$F(n) = Constant + AGF * (n)$$

F(n): forecast volume

AGF: average growth factor

n: the number of months from the first observation

Growth Factor

Methodology

Time Series Analysis

## Growth Factor

- Determination of constant and AGF from the linear regression
  - Tools / Add-ins / Analysis Tool Park
  - Tools / Data Analysis / Regression
  - Independent variable: month
  - Dependent variable: westbound truck volume (partial initial data)

| Month | Westbound truck volume |
|---|---|
| 1 | 13253 |
| 2 | 12878 |
| 3 | 14716 |
| 4 | 14186 |
| 5 | 15699 |
| 6 | 15619 |
| 7 | 13799 |
| 8 | 14612 |
| 9 | 14411 |
| 10 | 16232 |
| 11 | 15603 |
| 12 | 13525 |

Constant:13767

AGF: 119

Growth Factor

Methodology

Time Series Analysis

# Growth Factor Forecasting

F(n)=13767+119*n

| Month | Westbound truck volume | Months from first month | Linear regression |
|---|---|---|---|
| 1 | 13253 | 0 | 13767 |
| 2 | 12878 | 1 | 13886 |
| 3 | 14716 | 2 | 14005 |
| 4 | 14186 | 3 | 14124 |
| 5 | 15699 | 4 | 14243 |
| 6 | 15619 | 5 | 14362 |
| 7 | 13799 | 6 | 14481 |
| 8 | 14612 | 7 | 14600 |
| 9 | 14411 | 8 | 14719 |
| 10 | 16232 | 9 | 14838 |
| 11 | 15603 | 10 | 14957 |
| 12 | 13525 | 11 | 15076 |

Partial results with growth factor forecasting

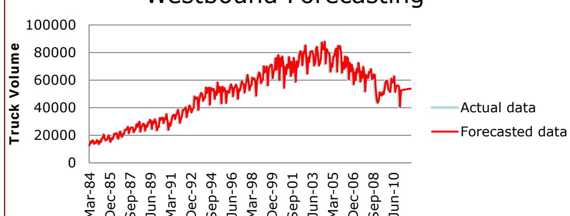
Growth Factor

Methodology

Time Series Analysis

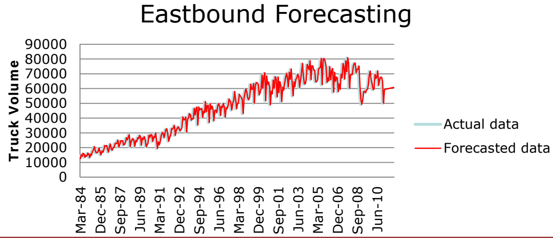
# Forecasting Results with All Initial Data

Westbound Forecasting




R Square 0.915

Eastbound Forecasting




R Square 0.919

Growth Factor

Methodology

Time Series Analysis

# Growth Factor Forecasting with Central Moving Average Series

F(n)=14291+230*n

Partial central moving average truck volume

| Month | Westbound truck volume | Months from first month | Linear regression |
|---|---|---|---|
| 1 | 14544 | 0 | 14291 |
| 2 | 14749 | 1 | 14410 |
| 3 | 14923 | 2 | 14529 |
| 4 | 15126 | 3 | 14648 |
| 5 | 15449 | 4 | 14767 |
| 6 | 15818 | 5 | 14886 |
| 7 | 15950 | 6 | 15005 |
| 8 | 16141 | 7 | 15124 |
| 9 | 16305 | 8 | 15243 |
| 10 | 16569 | 9 | 15362 |
| 11 | 16860 | 10 | 15481 |
| 12 | 17019 | 11 | 15600 |

Partial central moving average series

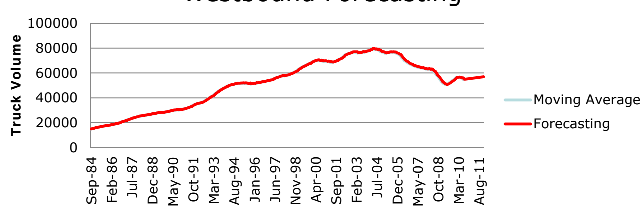
Growth Factor

Methodology

Time Series Analysis

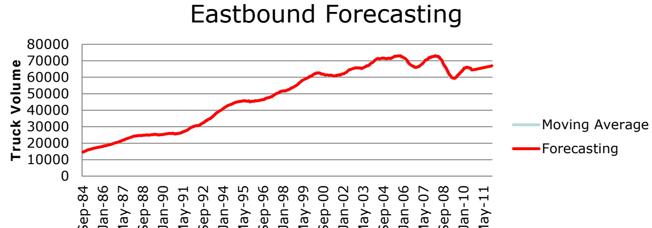
# Forecasting Results with All Smoothed Data

Westbound Forecasting




R Square 0.935

Eastbound Forecasting




R Square 0.940

Growth Factor

Methodology

Time Series Analysis

## Comparison

- Comparing R Square of Growth factor with initial data and moving average(smoothed) data

| R Square | Initial data | Moving average data |
|---|---|---|
| Westbound | 0.915 | 0.935 |
| Eastbound | 0.919 | 0.940 |

Growth Factor

Methodology

Time Series Analysis

## Growth Factor

- Compound growth

F(n)=Constant*AGF(n)

  - If there are two years

$$AGF = \left( F_2 / F_1 \right)^{1/(Y_2 - Y_1)}$$

Where $F_1$ is the freight flow in year $Y_1$, $F_2$ is the freight flow in year $Y_2$

  - If there are more than two years, AGF can be found from the linear regression

Growth Factor

Methodology

Time Series Analysis

# Growth Factor

- Determination of constant and AGF from the linear regression
  - o Tools / Add-ins / Analysis Tool Park
  - o Tools / Data Analysis / Regression
  - o Independent variable: month
  - o Dependent variable: westbound truck volume expressed as natural logarithm (partial initial data)
  - o Constant = EXP (intercept)
    AGF = EXP (x-variable coefficient)

| Date | Westbound | Ln |
|---|---|---|
| Jan-84 | 13253 | 9.4920 |
| Feb-84 | 12878 | 9.4633 |
| Mar-84 | 14716 | 9.5967 |
| Apr-84 | 14186 | 9.5600 |
| May-84 | 15699 | 9.6614 |
| Jun-84 | 15619 | 9.6562 |
| Jul-84 | 13799 | 9.5324 |
| Aug-84 | 14612 | 9.5896 |
| Sep-84 | 14411 | 9.5757 |
| Oct-84 | 16232 | 9.6947 |
| Nov-84 | 15603 | 9.6552 |
| Dec-84 | 13525 | 9.5123 |

Constant:13745
AGF: 1.008

Growth Factor

Methodology

Time Series Analysis

# Growth Factor Forecasting

$F(n) = 13745*1.008^n$

| Date | Westbound | Ln | Month from the first month | Compound regression |
|---|---|---|---|---|
| 1.00 | 13253 | 9.4920 | 0 | 13754 |
| 2.00 | 12878 | 9.4633 | 1 | 13864 |
| 3.00 | 14716 | 9.5967 | 2 | 13975 |
| 4.00 | 14186 | 9.5600 | 3 | 14087 |
| 5.00 | 15699 | 9.6614 | 4 | 14199 |
| 6.00 | 15619 | 9.6562 | 5 | 14313 |
| 7.00 | 13799 | 9.5324 | 6 | 14428 |
| 8.00 | 14612 | 9.5896 | 7 | 14543 |
| 9.00 | 14411 | 9.5757 | 8 | 14659 |
| 10.00 | 16232 | 9.6947 | 9 | 14777 |
| 11.00 | 15603 | 9.6552 | 10 | 14895 |
| 12.00 | 13525 | 9.5123 | 11 | 15014 |

Partial results with compound regression forecasting

Growth Factor

Methodology

Time Series Analysis

# Exponential Smoothing

- Model formulation:

$$S_1 = x_0$$
$$S_t = \alpha x_{t-1} + (1-\alpha) S_{t-1}, t > 1$$

where

$S_t$: exponentially smoothed value for time period t

$S_{t-1}$: exponentially smoothed value for time period t-1

$x_{t-1}$ : actual time series value for time period t

α: the *smoothing factor*, and 0 < α < 1

Exponential Smoothing

Time Series Analysis Methodology

# Example

$\alpha = 0.7$

$S_t = 0.7x_{t-1} + (1-0.7)S_{t-1}$

| x | S | Alpha |
|---|---|---|
| 13253 | | 0.7 |
| 12878 | $\boxed{13253}$ | |
| 14716 | $\boxed{12991}$ | |
| 14186 | 14198 | |
| 15699 | 14190 | |
| | 15246 | |

0.7*12878+0.3*13253

Exponential Smoothing

Time Series Analysis Methodology

# Exponential Smoothing

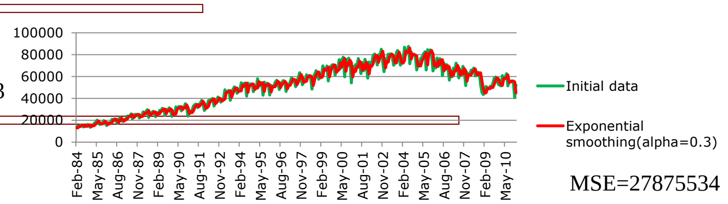
- Smoothing factor
  - The larger α is, the closer the smoothed value will track the original data value. The smaller α is, the more fluctuation is smoothed out.
- The determination of smoothing factor
  - Graph fitting
  - Mean squared error (MSE)
- Smoothing factor assumed (0.3, 0.5, 0.7)

Exponential Smoothing

Time Series Analysis

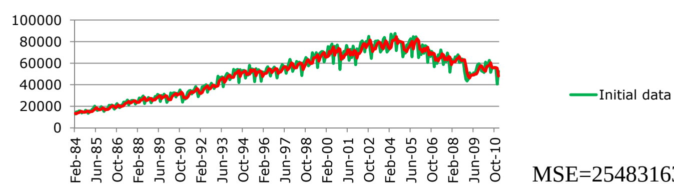
Methodology

Alpha=0.3



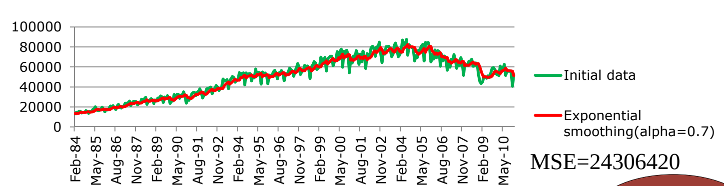

MSE=27875534

Alpha=0.5




MSE=25483163

Alpha=0.7




MSE=24306420

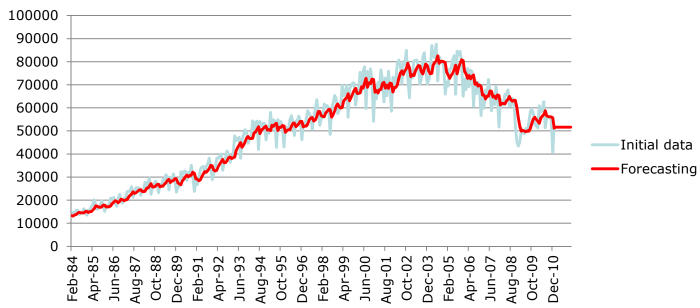
Time Series Analysis

Methodology

## Westbound Truck Volume Forecasting

| Alpha | 0.3 | 0.5 | 0.7 |
|---|---|---|---|
| MSE | 27875534 | 25483163 | 24306420 |

When alpha=0.7, MSE is smallest, so we select alpha=0.7 to forecast westbound truck volume




Exponential Smoothing

Time Series Analysis

Methodology

## Linear Regression

- Model

  $Y = f(x_1,x_2,\ldots,x_n) = b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n$

- Dataset with initial truck volume(partial):

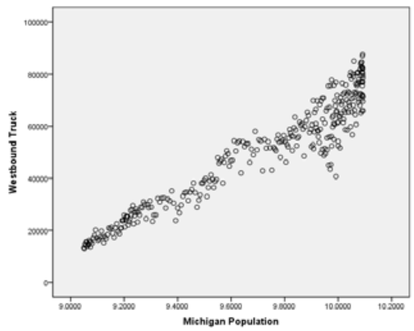
| Date | Moving Average Westbour | Eastbound | Population(million) Michigan | Ontario | US GDP(billion) | Fuel Price | NAFTA | Sep.11 |
|---|---|---|---|---|---|---|---|---|
| Jan-84 | 13253 | 12968 | 9.0490 | 8.9111 | 1269.133 | 1.216 | 0 | 0 |
| Feb-84 | 12878 | 12689 | 9.0513 | 8.9190 | 1269.133 | 1.209 | 0 | 0 |
| Mar-84 | 14716 | 14444 | 9.0535 | 8.9269 | 1269.133 | 1.210 | 0 | 0 |
| Apr-84 | 14186 | 13820 | 9.0558 | 8.9349 | 1302.100 | 1.227 | 0 | 0 |
| May-84 | 15699 | 15596 | 9.0580 | 8.9428 | 1302.100 | 1.236 | 0 | 0 |
| Jun-84 | 15619 | 15323 | 9.0603 | 8.9508 | 1302.100 | 1.229 | 0 | 0 |
| Jul-84 | 13799 | 13448 | 9.0625 | 8.9587 | 1325.333 | 1.212 | 0 | 0 |
| Aug-84 | 14612 | 14539 | 9.0648 | 8.9667 | 1325.333 | 1.196 | 0 | 0 |
| Sep-84 | 14411 | 14165 | 9.0670 | 8.9746 | 1325.333 | 1.203 | 0 | 0 |
| Oct-84 | 16232 | 15964 | 9.0693 | 8.9825 | 1344.667 | 1.209 | 0 | 0 |
| Nov-84 | 15603 | 15308 | 9.0715 | 8.9905 | 1344.667 | 1.207 | 0 | 0 |
| Dec-84 | 13525 | 13177 | 9.0738 | 8.9984 | 1344.667 | 1.193 | 0 | 0 |

Linear Regression

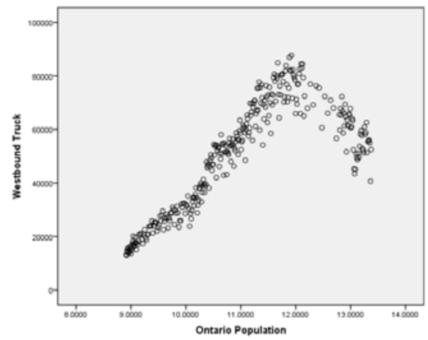
Time Series Analysis

Methodology

# Initial Analysis




Westbound truck & Michigan population scatterplot

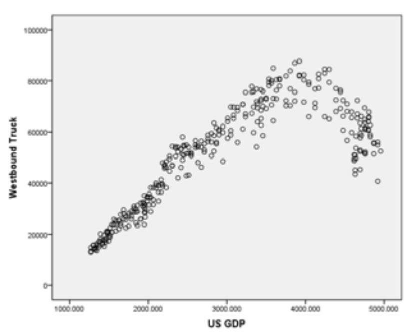



Westbound truck & Ontario population scatterplot

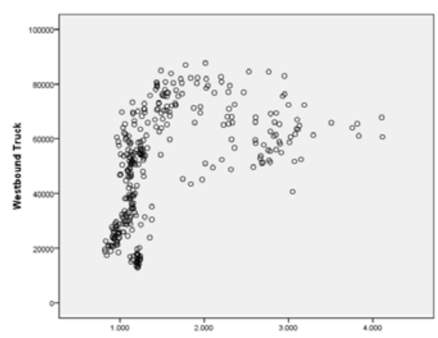
Time Series Analysis

Methodology

# Initial Analysis (cont'd)




Westbound truck & US GDP scatterplot

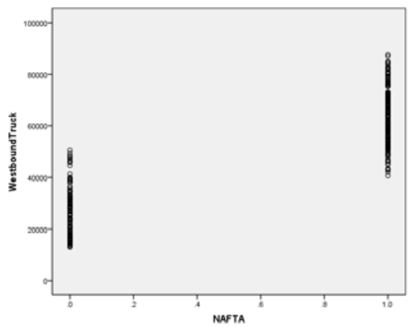



Westbound truck & Fuel price scatterplot

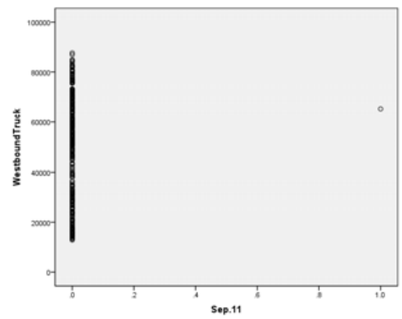
Time Series Analysis

Methodology

## Initial Analysis (cont'd)




Westbound truck & NAFTA scatterplot



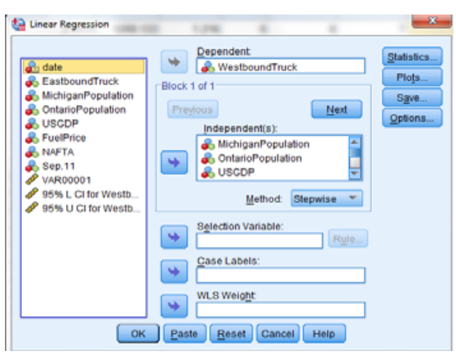

Westbound truck & September 11 scatterplot

Time Series Analysis

Methodology

## Regression Model Establishment with SPSS

- Select Analysis-Regression-Linear
- Put "Westbound Truck" as dependent variable and other variables as independent variables




- Select "Stepwise" as analysis mode
- Click OK

Linear Regression

Time Series Analysis

Methodology

# Results Analysis

**Variables Entered/Removed[a]**

| Model | Variables Entered | Variables Removed | Method |
|---|---|---|---|
| 1 | MichiganPopulation | . | Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100). |
| 2 | OntarioPopulation | . | Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100). |
| 3 | FuelPrice | . US GDP NAFTA, Sep.11 | Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100). |

a. Dependent Variable: WestboundTruck

The order of regression equation

The name of entered variables

The name of removed variables

The basis of entering and removing variables

Linear Regression

Methodology

Time Series Analysis

# Common Statistic

**Model Summary[d]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .953[a] | .908 | .907 | 6198.973 |
| 2 | .968[b] | .937 | .936 | 5141.600 |
| 3 | .971[c] | .942 | .942 | 4906.874 |

a. Predictors: (Constant), MichiganPopulation
b. Predictors: (Constant), MichiganPopulation, OntarioPopulation
c. Predictors: (Constant), MichiganPopulation, OntarioPopulation, FuelPrice
d. Dependent Variable: WestboundTruck

Adjusted R Square=0.942 indicates model 3 should be selected as regression model

Linear Regression

Methodology

Time Series Analysis

# Analysis of Coefficients

**Coefficients[a]**

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. |
|---|---|---|---|---|---|---|
| 1 | (Constant) | -490848.438 | 9612.915 | | -51.061 | .000 |
| | MichiganPopulation | 55777.692 | 990.233 | .953 | 56.328 | .000 |
| 2 | (Constant) | -618633.172 | 13201.094 | | -46.862 | .000 |
| | MichiganPopulation | 75922.463 | 1850.850 | 1.297 | 41.020 | .000 |
| | OntarioPopulation | -6145.543 | 505.998 | -.384 | -12.145 | .000 |
| 3 | (Constant) | -652201.511 | 13904.897 | | -46.904 | .000 |
| | MichiganPopulation | 83171.959 | 2175.991 | 1.421 | 38.223 | .000 |
| | OntarioPopulation | -10177.394 | 855.981 | -.636 | -11.890 | .000 |
| | FuelPrice | 4889.856 | 857.166 | .169 | 5.705 | .000 |

a. Dependent Variable: WestboundTruck

- Regression equation:

$$Y=-652201.511+83171.959x_1-10177.394x_2+4889.856x_3$$

$x_1$: Michigan population(million)

$x_2$: Ontario population(million)

$x_3$: Fuel price(current dollar)

Linear Regression

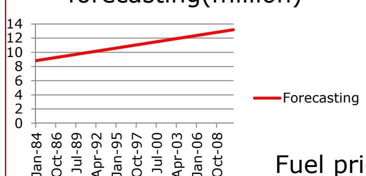
Time Series Analysis

Methodology

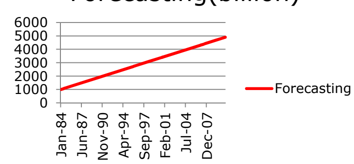
# Independent Variables Forecasting

- Forecasting the independent variables (Ontario population, US GDP, Fuel price ) with linear regression, respectively.

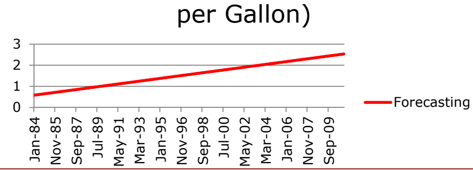
Ontario population forecasting(million)




US GDP Forecasting(billion)




Fuel price forecasting(Dollar per Gallon)




Linear Regression

Time Series Analysis

Methodology

# Michigan Population and GDP Forecasting

- Results of the 2010 headcount show the number of Michigan residents fell by 0.6 percent since 2000

| Year | Michigan population(million) |
|---|---|
| 2010 | 9.88 |
| 2011 | 9.29 |
| 2012 | 8.73 |
| 2013 | 8.21 |

=9.88*0.94

Forecasting the Michigan GDP

$$\text{Michigan GDP} = \left(\frac{\textit{Michigan Population}}{\textit{US Population}}\right) * \textit{US GDP}$$

Linear Regression

Time Series Analysis

Methodology

# Independent Variables Forecasting Results

- Independent variables forecasting results from Feb 2011 to Dec 2013

Partial forecasting results

| Date | Michigan | Ontario | US GDP(billion) | Fuel Price |
|---|---|---|---|---|
| Feb-11 | 9.241 | 12.978 | 4828.405 | 2.546 |
| Mar-11 | 9.194 | 12.991 | 4840.360 | 2.552 |
| Apr-11 | 9.148 | 13.004 | 4852.315 | 2.558 |
| May-11 | 9.101 | 13.017 | 4864.270 | 2.564 |
| Jun-11 | 9.055 | 13.03 | 4876.225 | 2.570 |
| Jul-11 | 9.009 | 13.043 | 4888.180 | 2.576 |
| Aug-11 | 8.962 | 13.056 | 4900.135 | 2.582 |
| Sep-11 | 8.916 | 13.069 | 4912.090 | 2.588 |
| Oct-11 | 8.869 | 13.082 | 4924.045 | 2.594 |
| Nov-11 | 8.823 | 13.095 | 4936 | 2.6 |
| Dec-11 | 8.776 | 13.108 | 4947.955 | 2.606 |

Linear Regression

Time Series Analysis

Methodology

# Regression Forecasting

- Westbound truck forecasting results from Feb 2011 to Dec 2013 with the multilinear regression model

| Date | Westbound | Michigan | Ontario | US GDP(billion) | Fuel Price |
|---|---|---|---|---|---|
| Feb-11 | 69696 | 9.241 | 12.978 | 4828.405 | 2.546 |
| Mar-11 | 69843 | 9.194 | 12.991 | 4840.360 | 2.552 |
| Apr-11 | 69989 | 9.148 | 13.004 | 4852.315 | 2.558 |
| May-11 | 70136 | 9.101 | 13.017 | 4864.270 | 2.564 |
| Jun-11 | 70282 | 9.055 | 13.03 | 4876.225 | 2.570 |
| Jul-11 | 70429 | 9.009 | 13.043 | 4888.180 | 2.576 |
| Aug-11 | 70575 | 8.962 | 13.056 | 4900.135 | 2.582 |
| Sep-11 | 70722 | 8.916 | 13.069 | 4912.090 | 2.588 |
| Oct-11 | 70868 | 8.869 | 13.082 | 4924.045 | 2.594 |
| Nov-11 | 71015 | 8.823 | 13.095 | 4936 | 2.6 |
| Dec-11 | 71162 | 8.776 | 13.108 | 4947.955 | 2.606 |

Partial forecasting results

Linear Regression

Time Series Analysis

Methodology

# Regression Model Establishment with EXCEL

- Inputting the dataset
- Click Anova Tools – Regression

| *Regression Statistics* | | |
|---|---|---|
| Multiple R | 0.971 | |
| R Square | 0.942 | Goodness of Fit >= 0.80 |
| Adjusted R Square | 0.942 | |
| Standard Error | 4906.954 | |
| Observations | 325 | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |
|---|---|---|---|---|
| Intercept | -652211.7462 | 13905.4354 | -46.90336745 | 0.000 |
| Michigan | 83173.84568 | 2176.085856 | 38.22176659 | 0.000 |
| Ontario | -10178.16051 | 856.0168203 | -11.89014079 | 0.000 |
| Fuel price | 4890.563481 | 857.1922136 | 5.705328867 | 0.000 |

y = -652211.746 +83173.846*Michigan -10178.161*Ontario +4890.563*Fuel Price

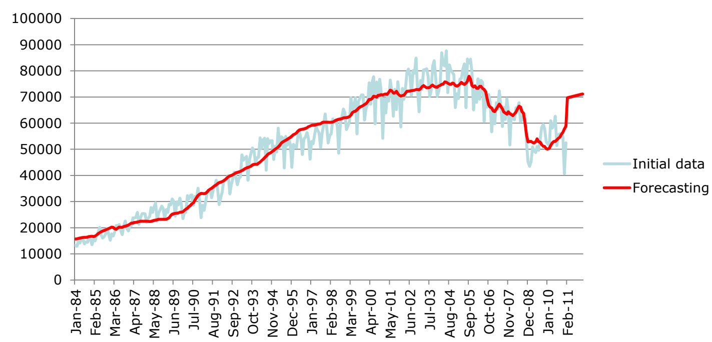
Linear Regression

Time Series Analysis

Methodology

# Forecasting Results

- Plot of initial value and forecasting value

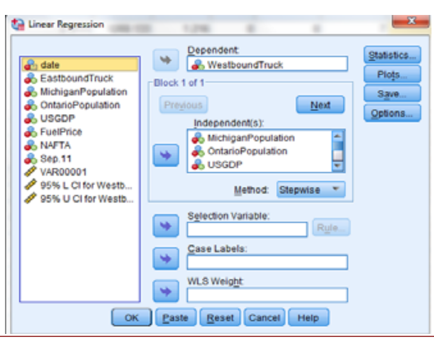



Linear Regression

Time Series Analysis Methodology

# Linear Regression with Smoothed Series(Central Moving Average Series)

- Select Analysis-Regression-Linear
- Put Westbound Truck as Dependent variable and other variables as independent variables




- Select Stepwise as analysis mode
- Click OK

Linear Regression

Time Series Analysis Methodology

# Regression Model Selection

**Model Summary**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .974[a] | .948 | .948 | 4452.129 |
| 2 | .990[b] | .979 | .979 | 2818.689 |
| 3 | .992[c] | .984 | .984 | 2450.851 |
| 4 | .992[d] | .985 | .984 | 2429.545 |

a. Predictors: (Constant), MichiganPopulation
b. Predictors: (Constant), MichiganPopulation, OntarioPopulation
c. Predictors: (Constant), MichiganPopulation, OntarioPopulation, USGDP
d. Predictors: (Constant), MichiganPopulation, OntarioPopulation, USGDP, NAFTA

Linear Regression

Time Series Analysis

Methodology

# Linear Regression Equation

**Coefficients[a]**

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. |
|---|---|---|---|---|---|---|
| 4 | (Constant) | -566514.597 | 13166.469 | | -43.027 | .000 |
| | MichiganPopulation | 76645.136 | 1293.081 | 1.335 | 59.273 | .000 |
| | OntarioPopulation | -13857.985 | 813.228 | -.890 | -17.041 | .000 |
| | USGDP | 9.095 | .911 | .526 | 9.982 | .000 |
| | NAFTA | -1733.426 | 677.358 | -.043 | -2.559 | .011 |

a. Dependent Variable: WestboundTruck

Y = -562356.809 +75631.264*Michigan -13266.173*Ontario +8.828*US GDP(billion) -1680.692*NAFTA

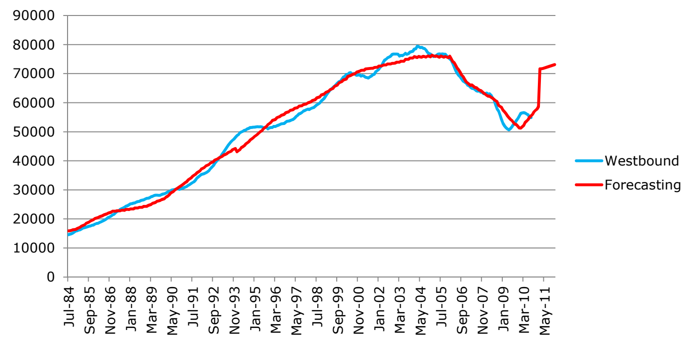
Linear Regression

Time Series Analysis

Methodology

# Forecasting with EXCEL

- Plot of westbound smoothed truck volume and forecasting volume




Linear Regression

Time Series Analysis Methodology

# Regression Forecasting with Seasonal Adjustment Factor

| Date | Forecasts | Seasonal adjustment factor | Adjusted forecasts |
|---|---|---|---|
| Sep-10 | 55,967 | 1.03 | 57,760 |
| Oct-10 | 56,876 | 1.08 | 61,596 |
| Nov-10 | 57,346 | 1.02 | 58,228 |
| Dec-10 | 57,816 | 0.86 | 49,863 |
| Jan-11 | 58,618 | 0.93 | 54,791 |
| Feb-11 | 71,656 | 0.91 | 65,562 |
| Mar-11 | 71,656 | 1.02 | 73,120 |
| Apr-11 | 71,816 | 1.00 | 71,694 |
| May-11 | 71,976 | 1.04 | 74,618 |
| Jun-11 | 72,136 | 1.04 | 74,697 |
| Jul-11 | 72,296 | 0.90 | 65,106 |
| Aug-11 | 72,456 | 1.03 | 74,838 |
| Sep-11 | 72,616 | 1.03 | 74,943 |
| Oct-11 | 72,776 | 1.08 | 78,816 |
| Nov-11 | 72,936 | 1.02 | 74,057 |
| Dec-11 | 73,096 | 0.86 | 63,042 |

Seasonal adjustment factor data is derived from central moving average series

Linear Regression

Time Series Analysis Methodology

# Comparison

- Comparing R Square of linear regression with smoothed data and unsmoothed data

| | R Square |
|---|---|
| Regression with unsmoothed data | 0.942 |
| Regression with smoothed data | 0.984 |

- Linear regression with smoothed data has a higher R square, this model can better fit the forecasted and actual value

Linear Regression

Time Series Analysis Methodology

# Linear Regression

- Michigan GDP as one of the independent variables instead of Michigan population and U.S. GDP

| Date | Westbound | Eastbound | Ontario population(million) | Fuel price | NAFTA | Sep.11 | Michigan GDP(billion) |
|---|---|---|---|---|---|---|---|
| Jan-84 | 13253 | 12968 | 8.9111 | 1.216 | 0 | 0 | 48.699 |
| Feb-84 | 12878 | 12689 | 8.9190 | 1.209 | 0 | 0 | 48.699 |
| Mar-84 | 14716 | 14444 | 8.9269 | 1.210 | 0 | 0 | 48.699 |
| Apr-84 | 14186 | 13820 | 8.9349 | 1.227 | 0 | 0 | 49.964 |
| May-84 | 15699 | 15596 | 8.9428 | 1.236 | 0 | 0 | 49.964 |
| Jun-84 | 15619 | 15323 | 8.9508 | 1.229 | 0 | 0 | 49.964 |
| Jul-84 | 13799 | 13448 | 8.9587 | 1.212 | 0 | 0 | 50.855 |
| Aug-84 | 14612 | 14539 | 8.9667 | 1.196 | 0 | 0 | 50.855 |
| Sep-84 | 14411 | 14165 | 8.9746 | 1.203 | 0 | 0 | 50.855 |
| Oct-84 | 16232 | 15964 | 8.9825 | 1.209 | 0 | 0 | 51.597 |
| Nov-84 | 15603 | 15308 | 8.9905 | 1.207 | 0 | 0 | 51.597 |
| Dec-84 | 13525 | 13177 | 8.9984 | 1.193 | 0 | 0 | 51.597 |

Time Series Analysis Methodology

# Forecasting

- Y=155753-20609*Ontario population(million) -13133*Fuel price+13033*NAFTA+1294*Michigan GDP(billion)

R Square: 0.989

| Date | Westbound | Ontario population | Fuel price | NAFTA | Michigan GDP |
|---|---|---|---|---|---|
| Feb-11 | 50039 | 12.978 | 2.546 | 1 | 140.76 |
| Mar-11 | 49076 | 12.991 | 2.552 | 1 | 140.29 |
| Apr-11 | 48108 | 13.004 | 2.558 | 1 | 139.81 |
| May-11 | 47138 | 13.017 | 2.564 | 1 | 139.32 |
| Jun-11 | 46163 | 13.03 | 2.57 | 1 | 138.84 |
| Jul-11 | 45186 | 13.043 | 2.576 | 1 | 138.35 |
| Aug-11 | 44204 | 13.056 | 2.582 | 1 | 137.86 |
| Sep-11 | 43220 | 13.069 | 2.588 | 1 | 137.37 |
| Oct-11 | 42232 | 13.082 | 2.594 | 1 | 136.87 |
| Nov-11 | 41240 | 13.095 | 2.6 | 1 | 136.38 |
| Dec-11 | 40245 | 13.108 | 2.606 | 1 | 135.87 |

Linear Regression

Time Series Analysis

Methodology

# Comparison

- Comparing R Square of linear regression with the independent variable of Michigan GDP and the independent variables without Michigan GDP

| | R Square |
|---|---|
| Regression with Michigan GDP | 0.984 |
| Regression without Michigan GDP | 0.989 |

- Linear regression with Michigan GDP has a higher R square, this model can better fit the forecasted and actual value

Time Series Analysis

Methodology

# ARIMA

- ARIMA(p,d,q)
  - Auto-regressive model
    p is the number of autoregressive terms
  - Integrated model
    d is the number of nonseasonal differences
  - Moving average model
    q is the number of lagged forecast errors in the prediction equation

ARIMA

Methodology

Time Series Analysis

# Auto-Regressive Model

- The definition of autoregressive(AR) model
  - Takes advantage of autocorrelation
    - 1st order - correlation between consecutive values
    - 2nd order - correlation between values 2 periods apart
- $p^{th}$ order Autoregressive models

$$Y_i = A_0 + A_1 Y_{i-1} + A_2 Y_{i-2} + \cdots + A_p Y_{i-p} + \delta_i$$

Random Error

ARIMA

Methodology

Time Series Analysis

# Autoregressive Modeling Example

- Develop the second order Autoregressive model

| Date | Westbound |
|---|---|
| Jan-84 | 13253 |
| Feb-84 | 12878 |
| Mar-84 | 14716 |
| Apr-84 | 14186 |
| May-84 | 15699 |
| Jun-84 | 15619 |
| Jul-84 | 13799 |
| Aug-84 | 14612 |

| Date | Yi | Yi-1 | Yi-2 |
|---|---|---|---|
| Jan-84 | 13253 | | |
| Feb-84 | 12878 | 13253 | |
| Mar-84 | 14716 | 12878 | 13253 |
| Apr-84 | 14186 | 14716 | 12878 |
| May-84 | 15699 | 14186 | 14716 |
| Jun-84 | 15619 | 15699 | 14186 |
| Jul-84 | 13799 | 15619 | 15699 |
| Aug-84 | 14612 | 13799 | 15619 |

Excel Output

| | Coefficients |
|---|---|
| Intercept | 13273 |
| X Variable 1 | -0.25 |
| X Variable 2 | 0.128 |

$$\hat{Y}_i = 13273 - 0.25Y_{i-1} + 0.128Y_{i-2}$$

Time Series Analysis

Methodology

# ACF and PACF

- ACF

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{covariance at lag k}}{\text{variance}}$$

$$= \frac{\sum (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum (Y_t - \bar{Y})^2}$$

$\rho_k$: The ACF at lag k.

| Lag | Autocorrelation |
|---|---|
| 1 | .227 |
| 2 | -.037 |
| 3 | -.122 |
| 4 | -.490 |
| 5 | -.067 |
| 6 | .028 |

| Lag | Partial Autocorrelation |
|---|---|
| 1 | .227 |
| 2 | -.094 |
| 3 | -.097 |
| 4 | -.474 |
| 5 | .171 |
| 6 | -.084 |

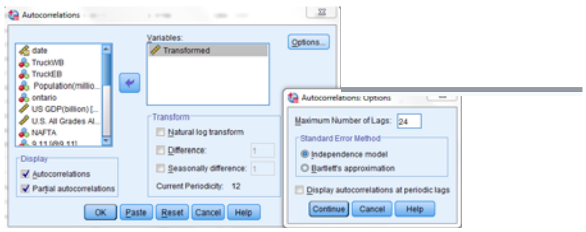
- PACF
  - The Partial Autocorrelation Function (PACF) is similar to the ACF, however it measures correlation between observations that are k time periods apart, after controlling for correlations at intermediate lags.

Time Series Analysis

Methodology

## Autocorrelation

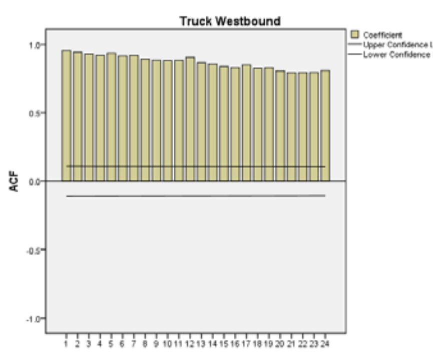
- Using initial truck series to make autocorrelation with SPSS
  - o Putting transformed data as variable
  - o Click "Autocorrelation" and "Partial autocorrelation" and input maximum number of lags as 24
  - o Click "OK"

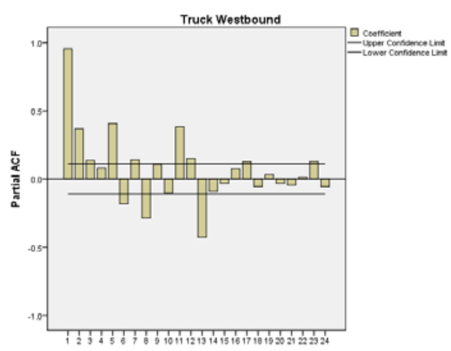



ARIMA

Time Series Analysis Methodology

## Autocorrelation of Initial Truck Series



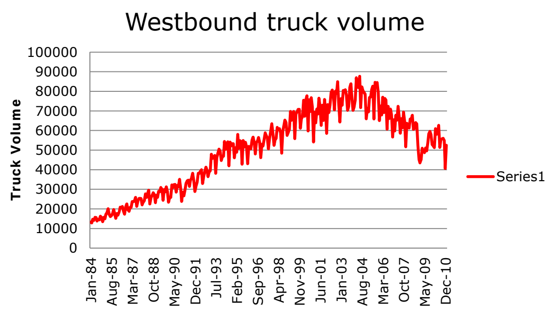




ARIMA

Time Series Analysis Methodology

# Identifying the Order of Differencing

- A stationary series
  - Constant mean
  - Constant variance
  - Constant autocorrelation structure




Nonstationary
The variability is changing

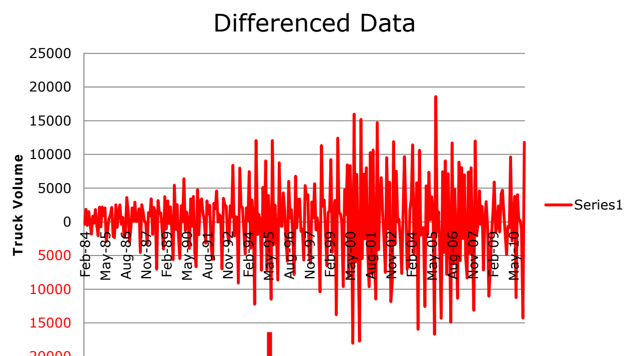
ARIMA

Time Series Analysis

Methodology

# Identifying the Order of Differencing

- Making the first differencing (d=1)




The series appears stationary, So we decide the d=1

ARIMA

Time Series Analysis

Methodology

## Removing the Changing of Variability

- Although the series appear stationary by differencing once, the variability is still changing over time. Transformation is considered for series in which variance changes over time and differencing does not stabilize the variance.

ARIMA



Methodology

## Box-Cox Transformation

- Transformation formulation

$$y_i^{(\lambda)} = \begin{cases} y_i^{\lambda}; & \lambda \neq 0 \\ \log y_i; & \lambda = 0 \end{cases}$$

Where $\lambda$ is transform parameter

ARIMA



Methodology

# Transform with EXCEL

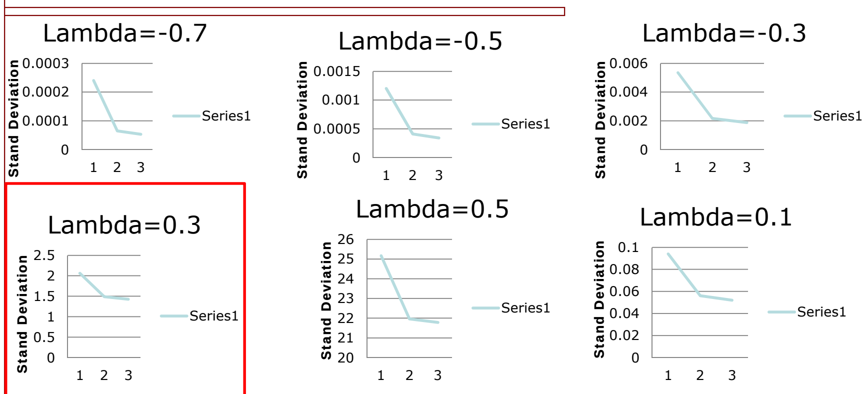
- Inputting the westbound truck volume (undifferenced) in EXCEL
- Click QI Macros – Anova Tools – Box cox
- Inputting the transform parameter lambda with -0.7, -0.5, -0.3, 0.1, 0.3, 0.5, respectively
- Dividing the transformed data into 3 components by time and computing the standard deviation of every component

| Order of components | Time interval |
| --- | --- |
| 1 | Jan 84-Apr 93 |
| 2 | May 93-Aug 02 |
| 3 | Sep 02-Jan 11 |

ARIMA

Time Series Analysis Methodology

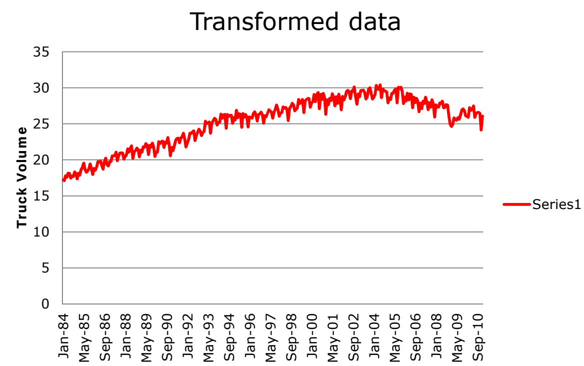
# Identifying Lambda




- When lambda is 0.3, the stand deviation of the three parts is most smoothing, we select lambda as 0.3.

Time Series Analysis Methodology

# Transforming with Undifferenced Data

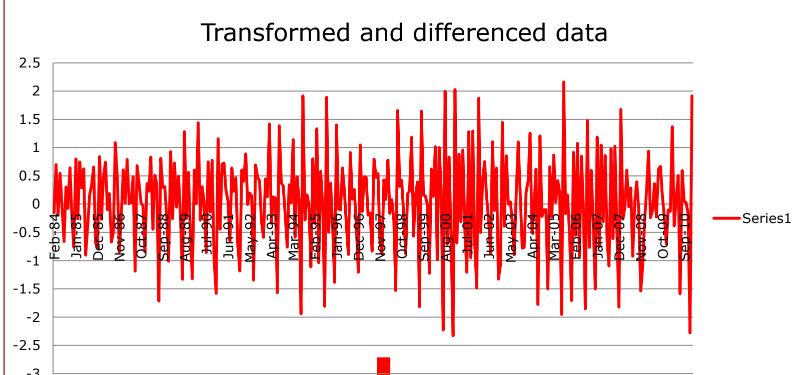



ARIMA

Methodology

Time Series Analysis

# Transforming with Differenced Data




Stationary
Seasonal

ARIMA

Methodology

Time Series Analysis

# Identifying AR(p)

- Identify the numbers of AR by looking at the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of differenced series

| | AR($p$) |
|---|---|
| ACF | Tails off |
| PACF | Cuts after $p$ |

ARIMA

Methodology

Time Series Analysis

# Moving Average Model

- The definition of moving average(MA) model

$$Y_t = \sum_{i=1}^{N} \theta_i a_{t-i} + a_t$$

$\theta_i$ : the correlation coefficient
$Y_t$ : the series under investigation
$a_t$ : the residual

- MA(2) model

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

ARIMA

Methodology

Time Series Analysis

## Identifying MA(q) Model

- Identify the numbers of MA by looking at the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of differenced series

| | MA($q$) |
|---|---|
| ACF | Cuts after $q$ |
| PACF | Tails off |

ARIMA

Methodology

Time Series Analysis

## Differencing=0

Series:Transformed

| Lag | Autocorrelation |
|---|---|
| 1 | .964 |
| 2 | .950 |
| 3 | .937 |
| 4 | .927 |
| 5 | .935 |
| 6 | .917 |
| 7 | .914 |
| 8 | .892 |
| 9 | .881 |
| 10 | .877 |
| 11 | .875 |
| 12 | .885 |
| 13 | .854 |
| 14 | .842 |
| 15 | .826 |
| 16 | .819 |
| 17 | .829 |
| 18 | .810 |
| 19 | .808 |
| 20 | .787 |
| 21 | .775 |
| 22 | .773 |
| 23 | .772 |
| 24 | .776 |

ACF falls to zero very slowly indicates that non-seasonal differencing is required

Series:Transformed

| Lag | Partial Autocorrelation |
|---|---|
| 1 | .964 |
| 2 | .302 |
| 3 | .107 |
| 4 | .070 |
| 5 | .336 |
| 6 | -.194 |
| 7 | .105 |
| 8 | -.259 |
| 9 | .101 |
| 10 | -.033 |
| 11 | .276 |
| 12 | .088 |
| 13 | -.385 |
| 14 | -.021 |
| 15 | -.027 |
| 16 | .120 |
| 17 | .122 |
| 18 | -.110 |
| 19 | .035 |
| 20 | -.046 |
| 21 | -.011 |
| 22 | .004 |
| 23 | .126 |
| 24 | -.059 |

ARIMA

Time Series Analysis

## Differencing=1

Series:Transformed

| Lag | Autocorrelation |
|---|---|
| 1 | -.402 |
| 2 | .008 |
| 3 | -.047 |
| 4 | -.322 |
| 5 | .466 |
| 6 | -.294 |
| 7 | .358 |
| 8 | -.213 |
| 9 | -.102 |
| 10 | -.044 |
| 11 | -.233 |
| 12 | .754 |
| 13 | -.351 |
| 14 | .097 |
| 15 | -.152 |
| 16 | -.283 |
| 17 | .489 |
| 18 | -.316 |
| 19 | .340 |
| 20 | -.132 |
| 21 | -.179 |
| 22 | -.020 |
| 23 | -.140 |
| 24 | .603 |

ACF cuts off to zero at lag 2 sharply, no non-seasonal differencing is needed

PACF falls to zero slowly and ACF cuts off to zero after lag 1,indicating MA(1) is needed

**Partial Autocorrelations**

Series:Transformed

| Lag | Partial Autocorrelation | Std. Error |
|---|---|---|
| 1 | -.402 | .056 |
| 2 | -.183 | .056 |
| 3 | -.144 | .056 |
| 4 | -.506 | .056 |
| 5 | .106 | .056 |
| 6 | -.224 | .056 |
| 7 | .288 | .056 |
| 8 | -.111 | .056 |
| 9 | .136 | .056 |
| 10 | -.509 | .056 |
| 11 | -.259 | .056 |
| 12 | .454 | .056 |
| 13 | .229 | .056 |
| 14 | .171 | .056 |
| 15 | -.050 | .056 |
| 16 | -.116 | .056 |
| 17 | -.028 | .056 |
| 18 | -.050 | .056 |
| 19 | -.027 | .056 |
| 20 | .124 | .056 |
| 21 | .073 | .056 |
| 22 | -.060 | .056 |
| 23 | .022 | .056 |
| 24 | .167 | .056 |

Time Series Analysis

## Identification of ARIMA(p,d,q)

P=0, d=1,q=1

The potential model is ARIMA(0,1,1)

$$(1-B)Y_t=(1-\theta_1 B)a_t$$

t: indexes time

B: the backshift operator, $BY_t= BY_{t-1}$

$\theta_1$: the nonseasonal moving average coefficient

$a_t$: the random error

ARIMA

Methodology

Time Series Analysis

## Model with Seasonal Components

- Seasonal model:

$$ARIMA(p,d,q)(P,D,Q)_S$$

where s is seasonal cycle
P: the order of seasonal AR model
D: the order of seasonal differencing
Q: the order of seasonal MA model

ARIMA
Methodology



## Identification of P,D,Q

Differencing: 0

**Autocorrelations**

Series:Transformed

| Lag | Autocorrelation | Std. Error[a] | Box-Ljung Statistic Value | df | Sig.[b] |
|---|---|---|---|---|---|
| 12 | .754 | .054 | 461.271 | 12 | .000 |
| 24 | .603 | .053 | 850.763 | 24 | .000 |

a. The underlying process assumed is independence (white noise).
b. Based on the asymptotic chi-square approximation.

**Partial Autocorrelations**

Series:Transformed

| Lag | Partial Autocorrelation | Std. Error |
|---|---|---|
| 12 | .454 | .056 |
| 24 | .167 | .056 |

ACF falls to zero slowly at lags 12 and 24, indicating seasonal differencing is needed

ARIMA
Methodology

# ACF and PACF of Seasonal Differenced Series

Differencing : 1

**Autocorrelations**

Series:Transformed

| Lag | Autocorrelation | Std. Error[a] | Box-Ljung Statistic Value | df | Sig.[b] |
|---|---|---|---|---|---|
| 12 | -.116 | .055 | 171.673 | 12 | .000 |
| 24 | -.281 | .054 | 335.893 | 24 | .000 |

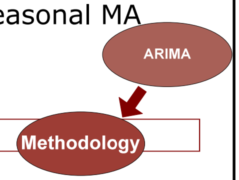
a. The underlying process assumed is independence (white noise).
b. Based on the asymptotic chi-square approximation.

**Partial Autocorrelations**

Series:Transformed

| Lag | Partial Autocorrelation | Std. Error |
|---|---|---|
| 12 | -.219 | .057 |
| 24 | -.102 | .057 |

- ACF cuts off to zero at lags 12 and 24 indicates no more seasonal differencing is needed.
- Both ACF and PACF do not show to cut off to zero at lags 12 and 24, indicating no seasonal AR model and seasonal MA model are needed.




Time Series Analysis

# Identification of Final ARIMA(p,d,q)(P,D,Q)$_S$

p=0,d=1,q=1

P=0,D=1,Q=0

ARIMA(0,1,1)(0,1,0)$_{12}$

$$(1-B)(1-B^{12})Y_t=(1-\theta_1 B)a_t$$

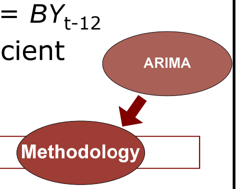
t: indexes time

B: the backshift operator, $BY_t = BY_{t-1}$

$B^{12}$: the seasonal backshift operator, $B^{12}Y_t = BY_{t-12}$

$\theta_1$: the nonseasonal moving average coefficient

$a_t$: the random error




Time Series Analysis

# Model Building with SPSS

**Model Description**

| | | | Model Type |
|---|---|---|---|
| Model ID | Transformed | Model_1 | ARIMA(0,1,1)(0,1,0) |

**Model Statistics**

| Model | Number of Predictors | Model Fit statistics | | Ljung-Box Q(18) | | | Number of Outliers |
|---|---|---|---|---|---|---|---|
| | | Stationary R-squared | R-squared | Statistics | DF | Sig. | |
| Transformed-Model_1 | 1 | 8.104E-6 | .971 | 232.645 | 18 | .000 | 0 |

R square : 0.971

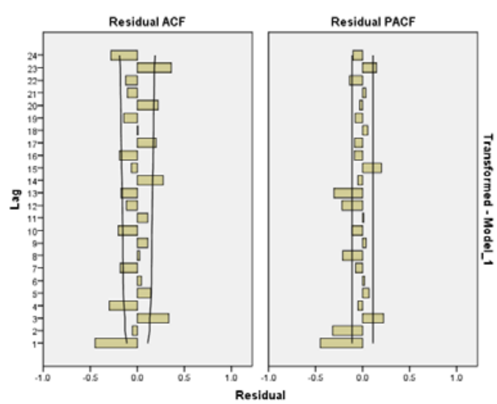
**ARIMA Model Parameters**

| | | | | | Estimate | SE | t | Sig. |
|---|---|---|---|---|---|---|---|---|
| Transformed-Model_1 | Transformed | No Transformation | Constant | | -.005 | .014 | -.326 | .745 |
| | | | Difference | | 1 | | | |
| | | | MA | Lag 1 | .482 | .050 | 9.685 | .000 |
| | | | Seasonal Difference | | 1 | | | |

Time Series Analysis

Methodology

# Residual Graph

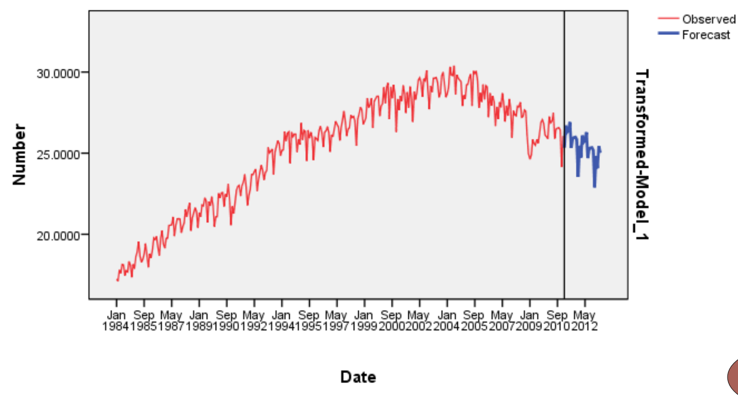



ARIMA

Methodology

Time Series Analysis

# Forecasting




ARIMA

Time Series Analysis | Methodology

# Forecasting Results (Partial)

| Date | Westbound Truck Volume | Transformed forecasts |
|---|---|---|
| Feb-11 | 51049 | 25.9304 |
| Mar-11 | 61259 | 27.2993 |
| Apr-11 | 57799 | 26.9157 |
| May-11 | 58566 | 27.0224 |
| Jun-11 | 62301 | 27.5289 |
| Jul-11 | 51151 | 25.9460 |
| Aug-11 | 55153 | 26.5395 |
| Sep-11 | 55538 | 26.5950 |
| Oct-11 | 55720 | 26.6213 |
| Nov-11 | 54643 | 26.4657 |
| Dec-11 | 40496 | 24.1885 |

Forecasts of truck volume derived from the forecasts of transformed series

Forecasts of transformed series

ARIMA

Time Series Analysis | Methodology

## Conclusions

- Growth factor
- Exponential smoothing
- Linear regression
- ARIMA




## Conclusions (cont'd)

- Growth factor
  - o Advantages
    - Simple and easy to understand
    - Considering the linear and nonlinear trend of the historical data
  - o Disadvantages
    - Neglecting the effects of cyclical or seasonal components
    - Increasing the time

## Conclusions (cont'd)

- Exponential smoothing
  - o Advantages
    - Including both linear and nonlinear
    - Structural view of the data that include level, trend, seasonality, and events
  - o Disadvantage
    - The forecast is constant for all future values



## Conclusions (cont'd)

- Linear regression
  - o Advantage
    - Considering the characteristic of independent variables and the relationship between independent variables(economical and political factors) and dependent variables
  - o Disadvantage
    - Difficult to determine the trends of every independent variable

## Conclusions (cont'd)

- ARIMA
  - Advantage
    - The comprehensiveness of the family of models
  - Disadvantages
    - ARIMA identification is difficult and time consuming
    - ARIMA may be difficult to explain to others
    - Models that perform similarly on the historical data may yield quite different forecasts
    - Empirical
  - Although there are more disadvantages than advantages, the advantages may still outweigh the disadvantages.



## Reference

1. Daniel Beagan, Michael Fischer, Arun Kuppam. "Quick Response Freight Manual II" (2007).
2. Alan Pankratz. "Forecasting With Univariate Box-Jenkins Models" (1983).
3. R.M.SAKIA. "The Box-Cox transformation technique: a review" (1992).