

Forecast and Capacity Planning for Nogales' Ports of Entry

Research Team:

Dr. Rene Villalobos, Dr. Arnold Maltz, Liangjie Xue,
Octavio Sanchez and Laura Vasquez

Arizona State University



December 2009



Agenda

- Objective of the study
- Executive summary
- Data collection and discussion
 - Data Review
 - Variable Selection
- Model Selection
 - Model types
 - Model selection
- Forecast and discussion
 - Updated Models
 - Scenarios discussion
- Simulation Results
- Conclusion

Objective of the Study

- The principal objective of this study is to forecast the number of border crossings by mode at the Nogales-Mariposa and DeConcini Ports of Entry (POEs)
- A secondary objective is the assessment of the interaction between the Mariposa and DeConcini Ports of Entry
- A third objective is the assessment of future port of entry needs and opportunities

Executive Summary

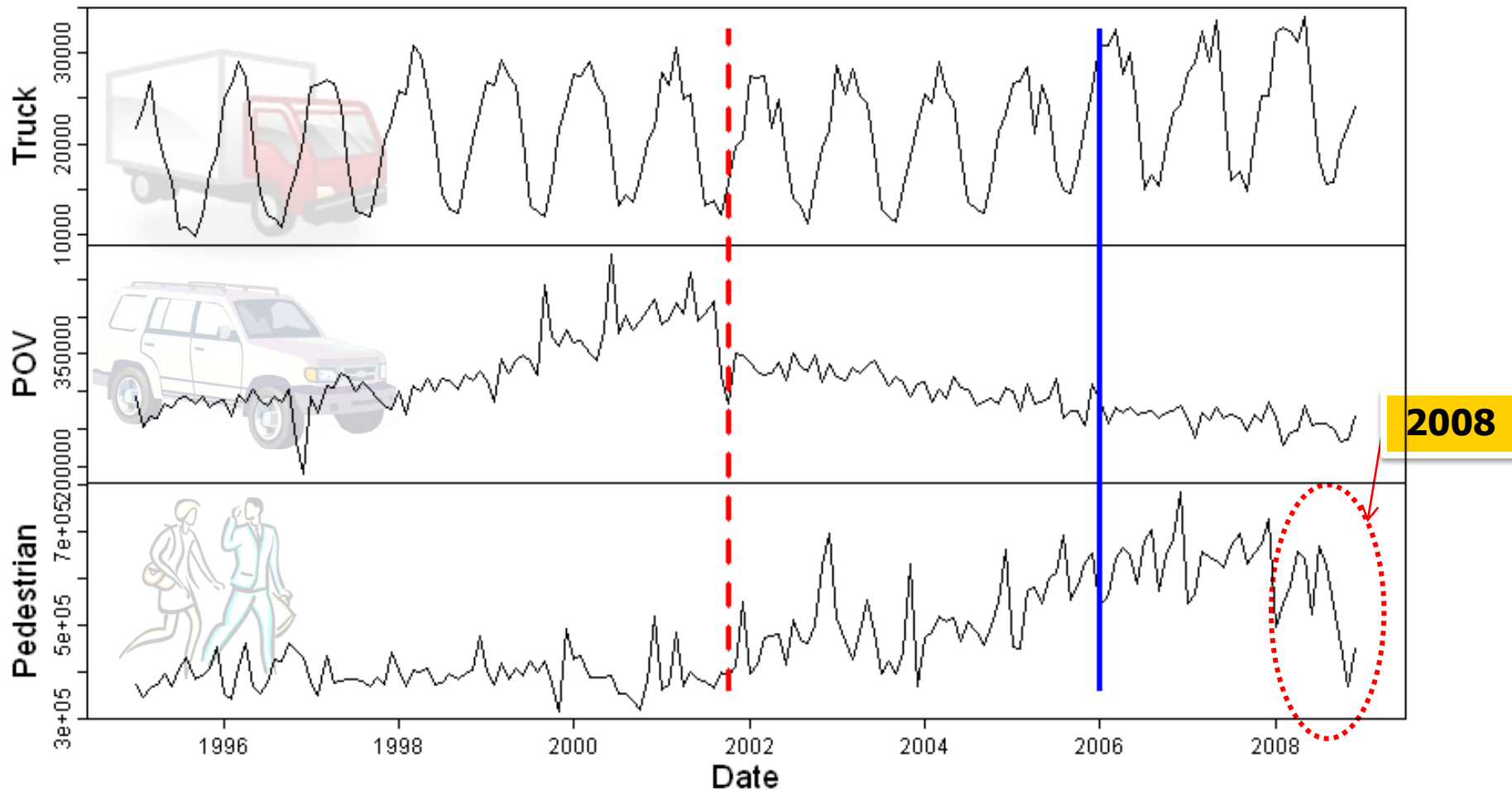
- We tested and used both time series and regression models to prepare 5, 10, and 15 year forecasts
- The Mexican Peso to US Dollar exchange rate and the US Index of Industrial Production were the only external drivers of cross-border traffic that surfaced in the research
- Truck crossings may increase by 50% vs. 2008 in the next 15 years, although the recent recession may delay this growth
- Privately owned vehicle (POV) and pedestrian traffic is also likely to increase, but is much more sensitive to specific economic events and thus harder to project.
- Bus passenger traffic remains a small portion of overall crossings

DATA COLLECTION AND MODEL SELECTION

- Data Review
- Variable Selection

Historical data for the primary modes

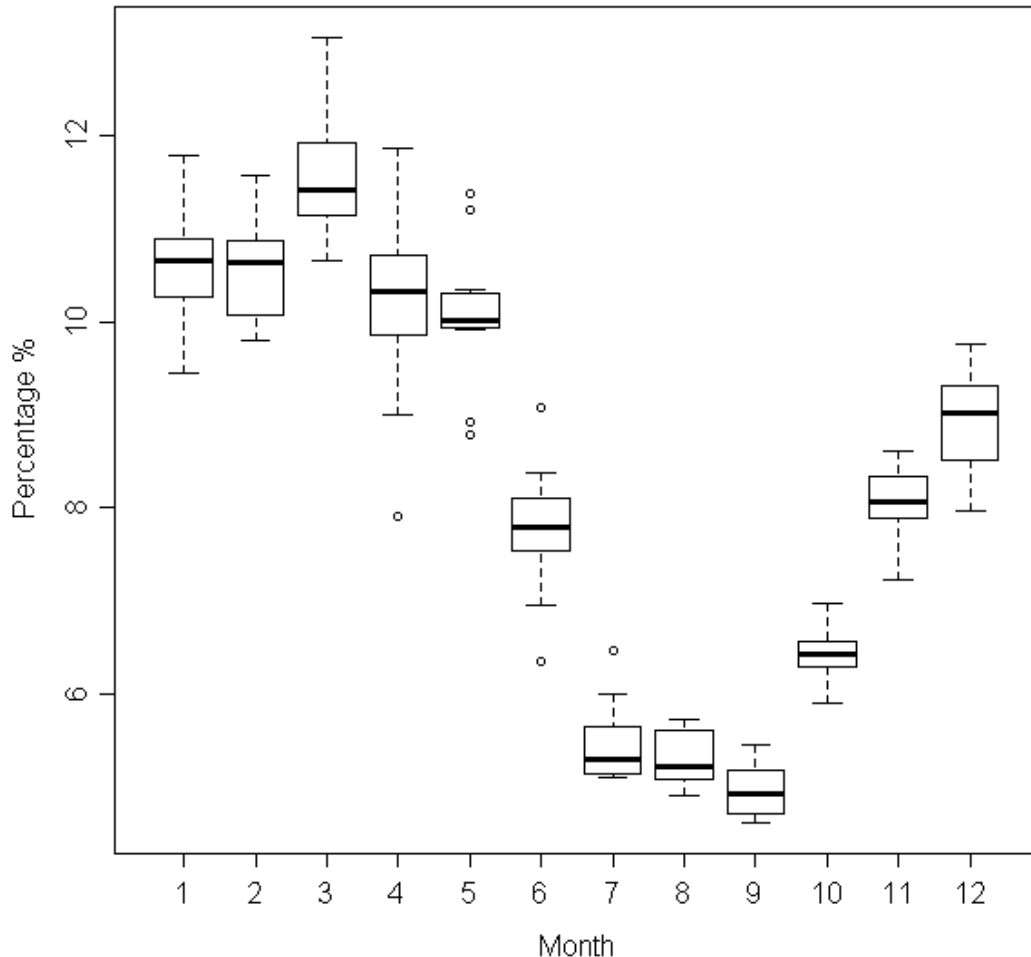
Border Crossing of each mod at Nogales (01/1995 to 12/2008)



* 9/11

* Testing Data

The seasonality in the truck traffic



- Based on 14 years of history, we identified a fairly stable seasonal pattern. This quantifies the effects of Nogales' position in the produce supply chain.
- The stability of the pattern allowed us to disaggregate yearly results as necessary.

Variables investigated for causal modeling

Data Name	Time range	Frequency
US national GDP	from 1949 to 2008 Q4	quarterly
Mexican national GDP	from 1993 to 2008 Q4	quarterly
Exchange rate (1USD in MNX)	Since Jan 1994	daily, monthly
Arizona GDP	1997 -2007	yearly
US fuel price (Gasoline and Diesel)	Jan 1994 to Dec 2008	Monthly
Arizona Population	1990 to 2008	yearly
Sonora Population	1995 to 2008	yearly
US Index of Industrial Production (IIP)	Since 1919	monthly
MX Index of Industrial Production (IIP)	Since 1990	monthly
US Consumer Price index (CPI)	Since 1990	monthly
MX Consumer Price index (CPI)	Since 1990	Monthly
Real exchange rate	Calculated from exchange rate and CPIs Since Jan 1994	monthly

Method Overview

- Modeling Process
 - Final three years of data used to test models derived from history before that
 - Evaluate all the possible combinations of candidate variables up to 5 variables in the model to be tested
- Selection Criteria:
 - Theil's U statistic (The smaller the better)
 - R-square value (The bigger the better)
 - VIF value of the variables (usually should be below 10)
 - Practical meaning of the model
- Details in the Appendix

Model Alternatives

- Regression Models
 - Multivariate model
- Time Series Models
 - Univariate, consider a method named "Holt-Winter's Method"
 - Multivariate: including exogenous variables
 - Considered ARIMA model, a category of time series model
 - Same type of models with different parameters have different performances

Model Selection: Example

- Regression Model Coefficients:

Intercept	2.984 e-16
------------------	------------

USIIP	5.545 e-01
--------------	------------

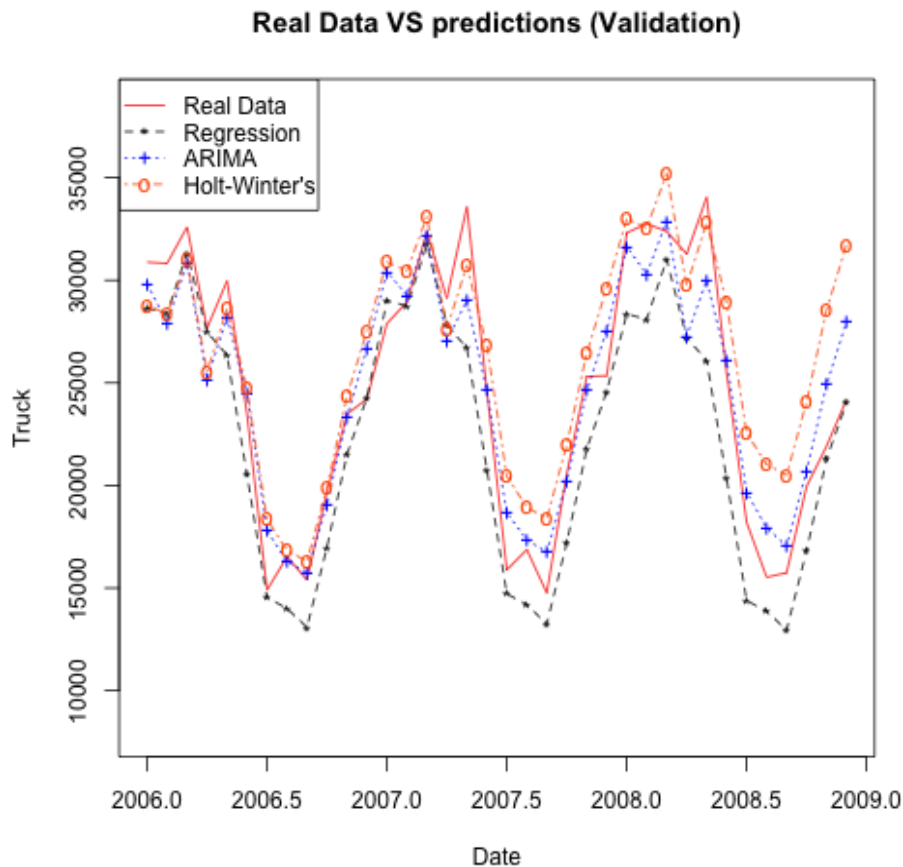
X-Rate	5.529 e-01
---------------	------------

- ARIMA parameters

$$(p,d,q)(P,D,Q)_L =$$

$$(1,1,4)(2,1,2)_{12}$$

*L is the seasonal period



Model Selection: Example

Performance comparison on Validation set

Method	R square (The higher the better)	Theil's U statistic (The lower the better)
Multivariate Regression	0.765	0.06315865
Holt-Winter's	0.760	0.05936151
Multivariate time series	0.889	0.04156882

- Time series model outperforms the regression model in terms of both criteria, hence the time series model was adopted

FORECAST AND DISCUSSION

Finalizing the Model

- The actual forecast model was based on the full data set, thus including the latest three years of data which were previously used for model testing
- In general, adding the latest three years did not change the model structure or parameter selection
- Given the relatively long time horizons, we used multiple scenarios to test the levels of uncertainty in the forecasts

Truck Crossings: Forecast



Commercial: Forecast Overview

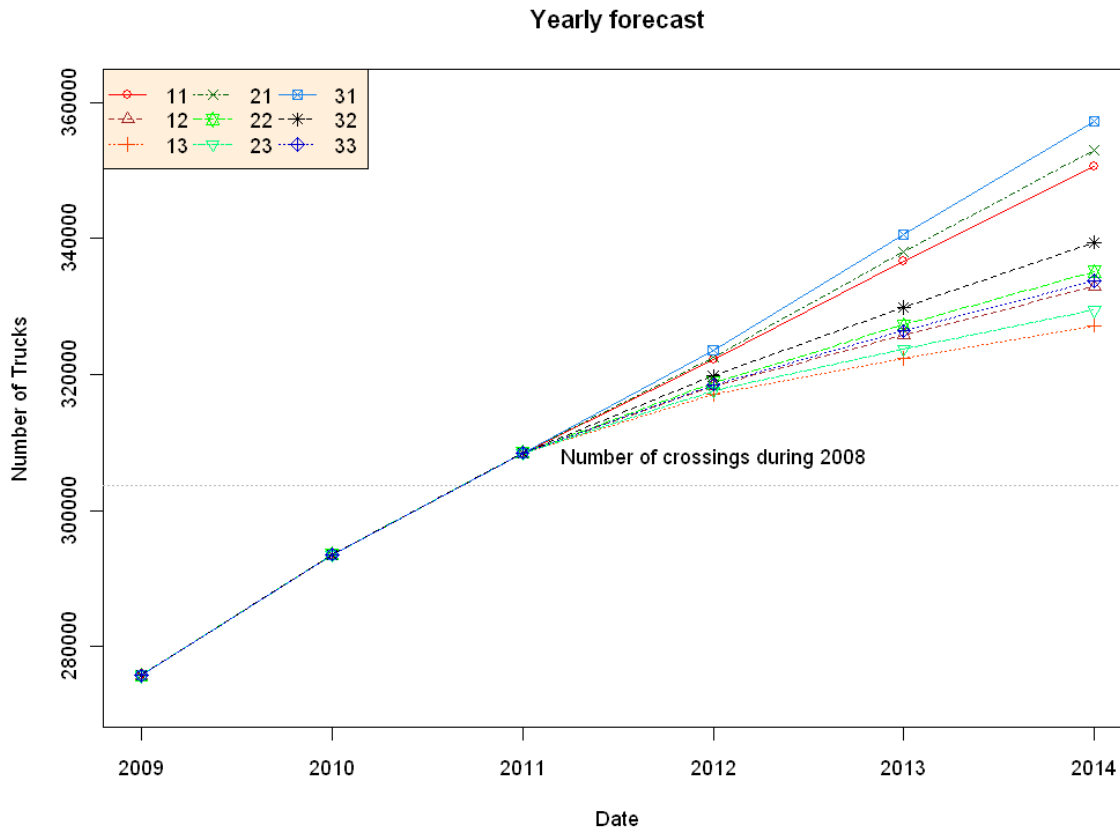
- External variables:
 - Mexican Peso to US Dollar Exchange Rate
 - US Index of Industrial Production (IIP)
- For each mode of traffic we provided five-year, ten-year and fifteen-year forecasts
- We used exchange rate and US IIP forecasts from forecasts.org, for the initial 3 years and
- Created different scenarios for these external variables beyond 36 month time frame

Scenarios

- Total of 9 scenarios for Exchange Rate and US IIP combinations
- Details available from project team
- Varying levels of data available to support medium and long term forecasts

Level	Exchange Rate	US IIP
1	Growing Fast	Growing Fast
2	Growing Mildly	Growing Mildly
3	Staying relatively stable	Staying relatively stable

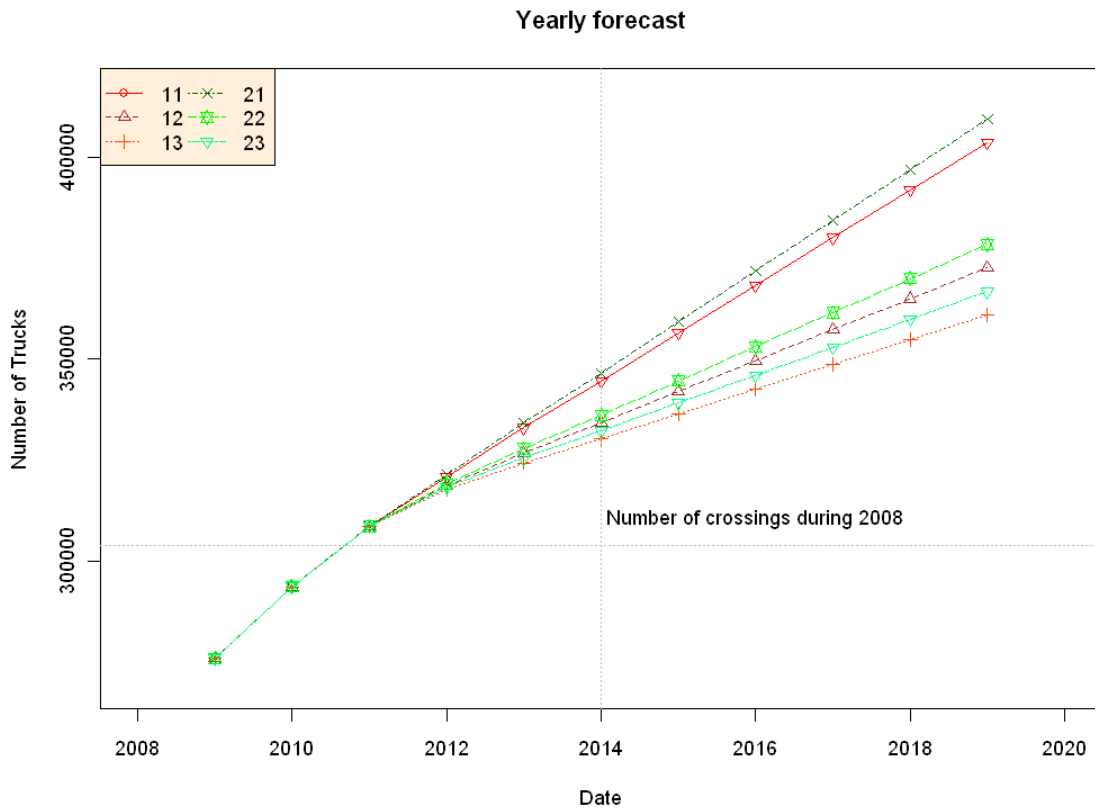
Truck: 5-year Forecast



**Change by 2014
(%) 2008=100**

		X-Rate		
		+		-
US IIP	+	1/1	2/1	3/1
		15.4	16.2	17.6
	+	1/2	2/2	3/2
		9.6	10.3	11.7
	-	1/3	2/3	3/3
		7.7	8.5	9.9

Truck: 10-year Forecast

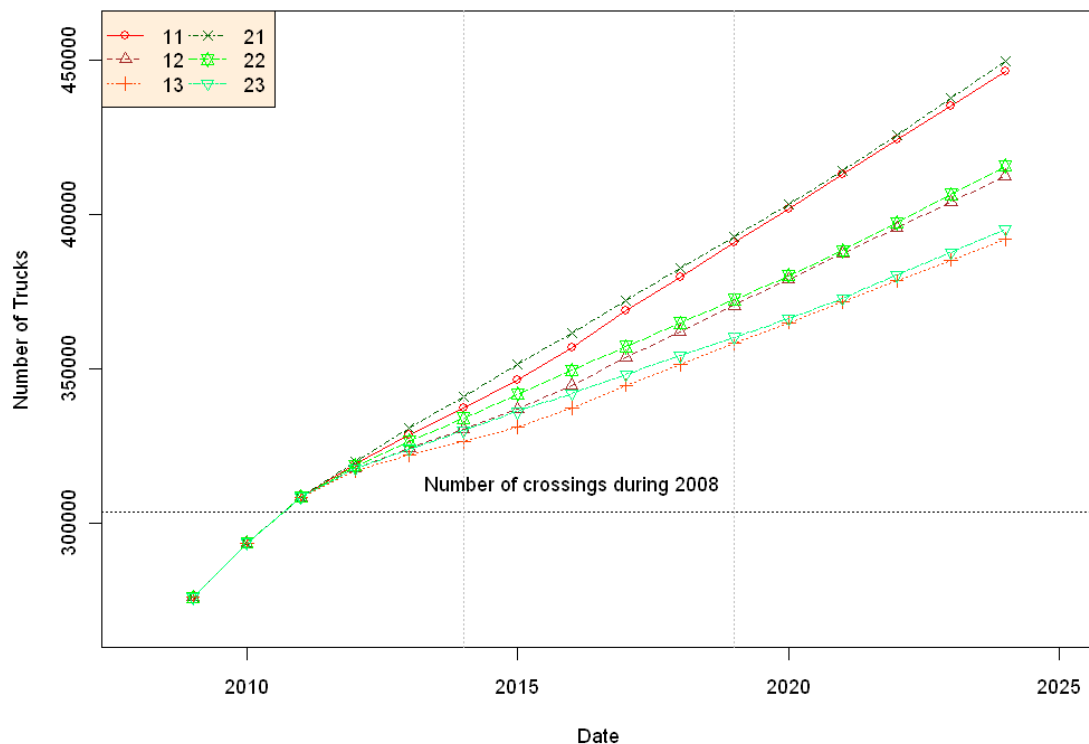


Change by 2019 (%)
2008=100

		X-Rate	
		+	-
USIIP	+	Growth Speed	
	+	1/1	2/1
		32.9	34.8
	+	1/2	2/2
		22.7	24.6
	+	1/3	2/3
-	18.8	20.8	

Truck: 15-year Forecast

Yearly forecast



**Change by 2024
(%) 2008=100**

X-Rate

+ Growth Speed -

USIIP +

-

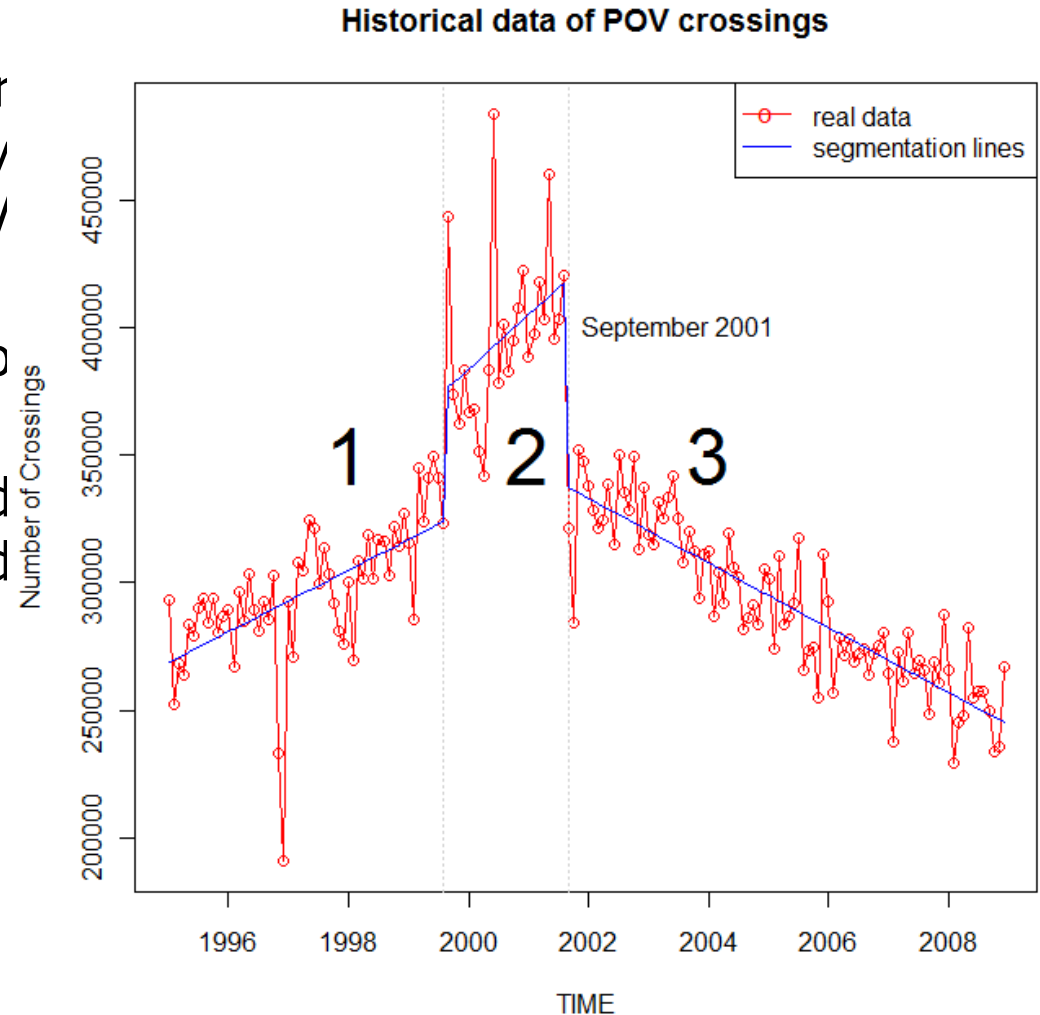
1/1	2/1
47.2	42.3
1/2	2/2
35.9	37.0
1/3	2/3
29.1	30.2

Privately Owned Vehicles: Forecast



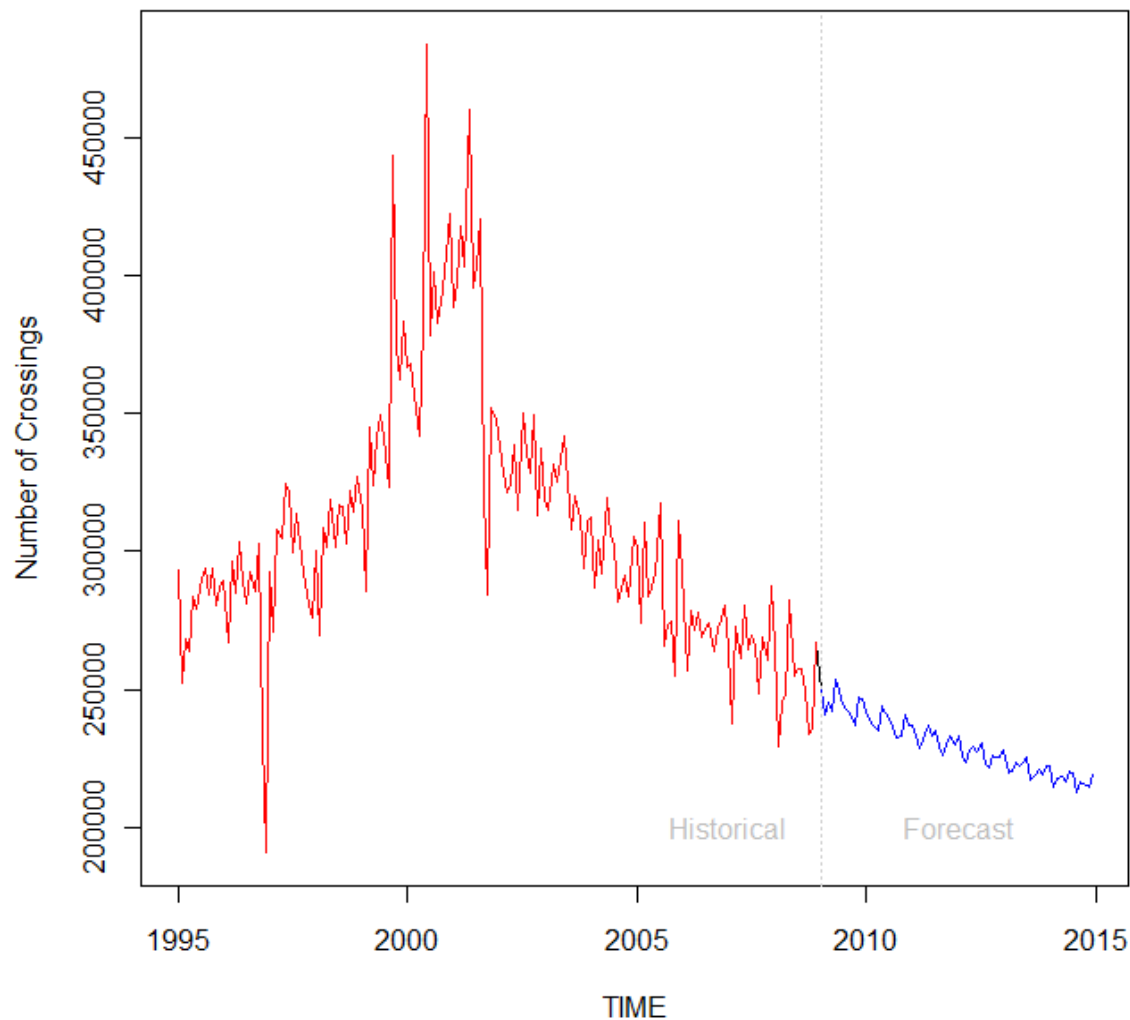
POV: Forecast Overview

- There was no external factor that was statistically significant to the POV crossings
- ARIMA model was used to forecast the 5-year trend
- A simple regression method was used for the extended forecast



POV: 5-year Forecast

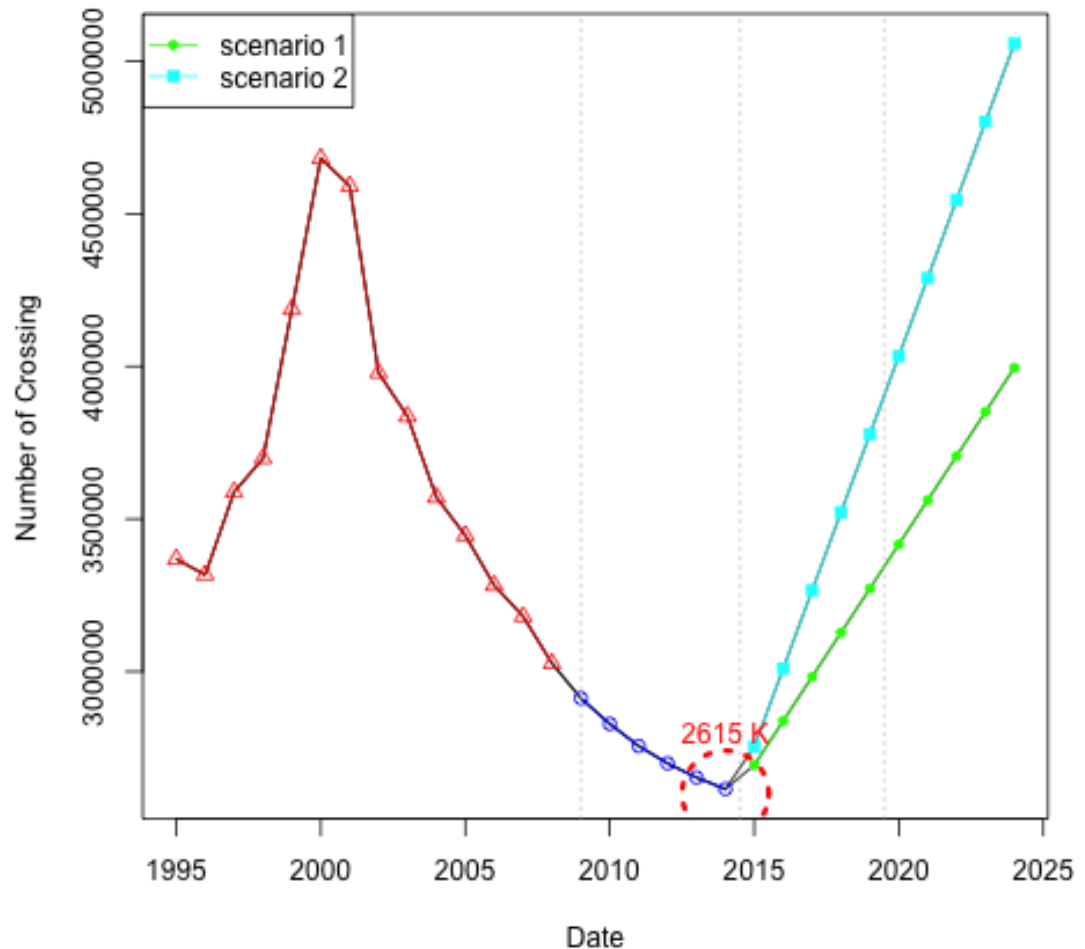
5-Year Forecast of the POV Crossing



POV: 10-year & 15-year Forecast

- We assumed the crossing traffic would start to increase after the current recession is over (red dashed circle)
- Recession bottom for crossing purposes at 2014 simply to identify a starting point for growth

10&15-Year Forecast of the POV Crossing

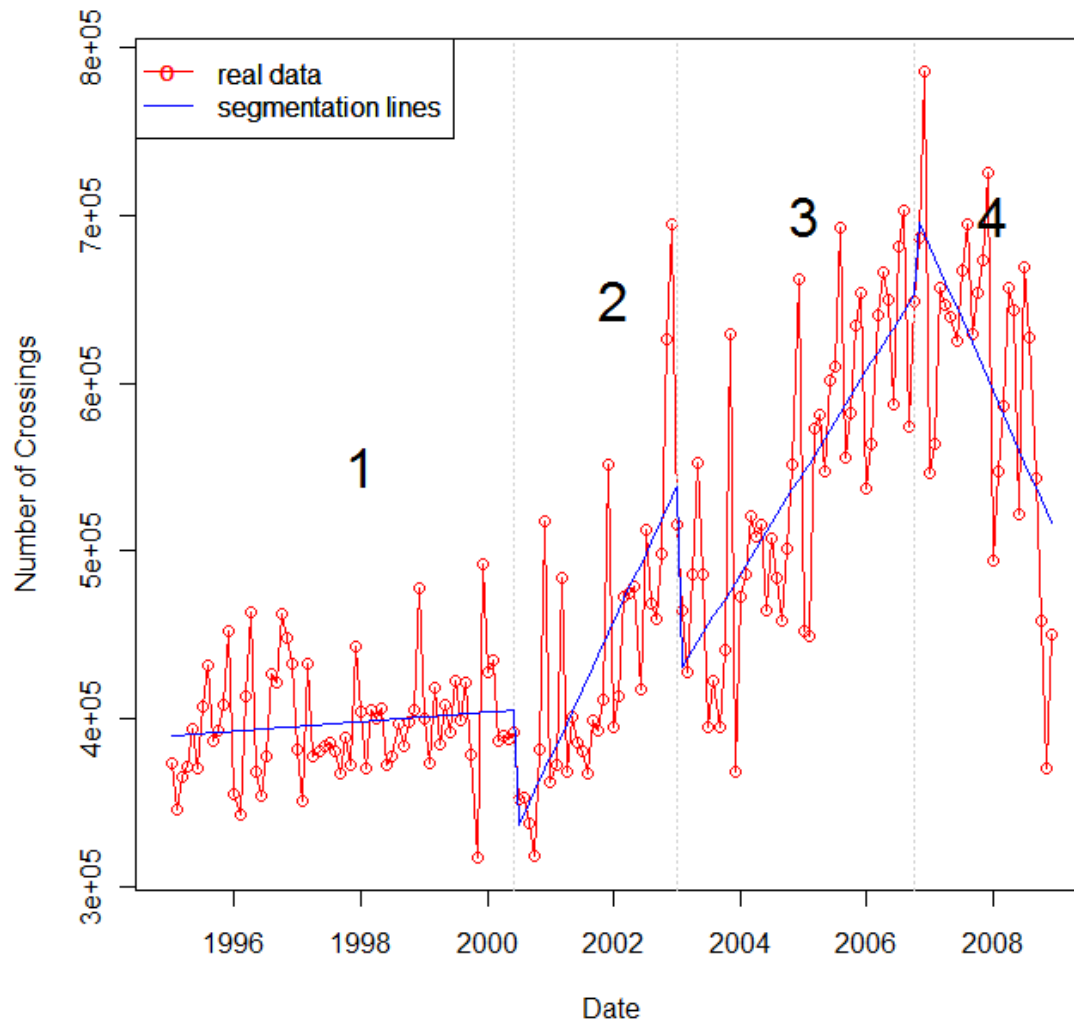


Pedestrian: Forecasts



Pedestrian: Historical Data Review

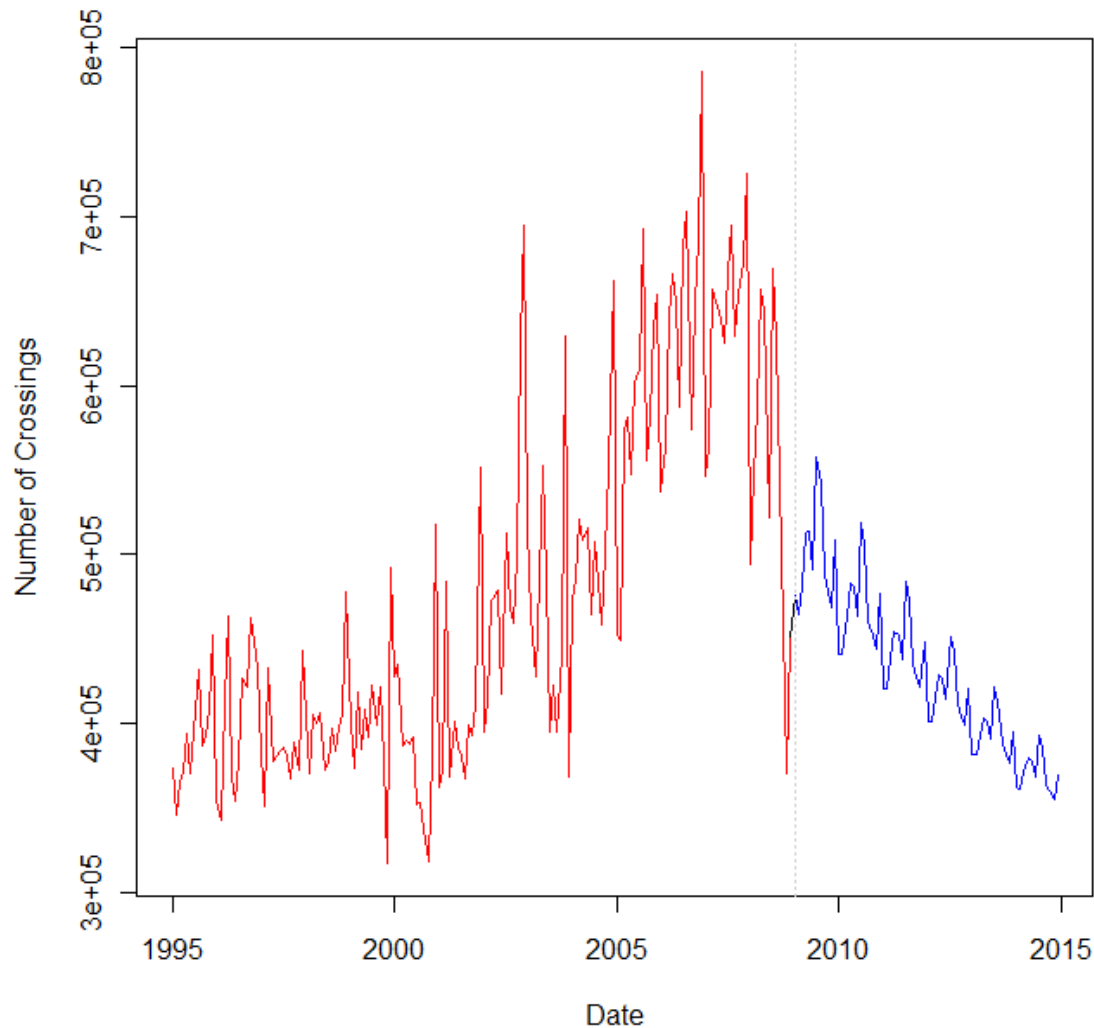
Historical data of Pedestrian crossings



- The historical data can be divided into four different segments
- Each segment has a different increasing trend

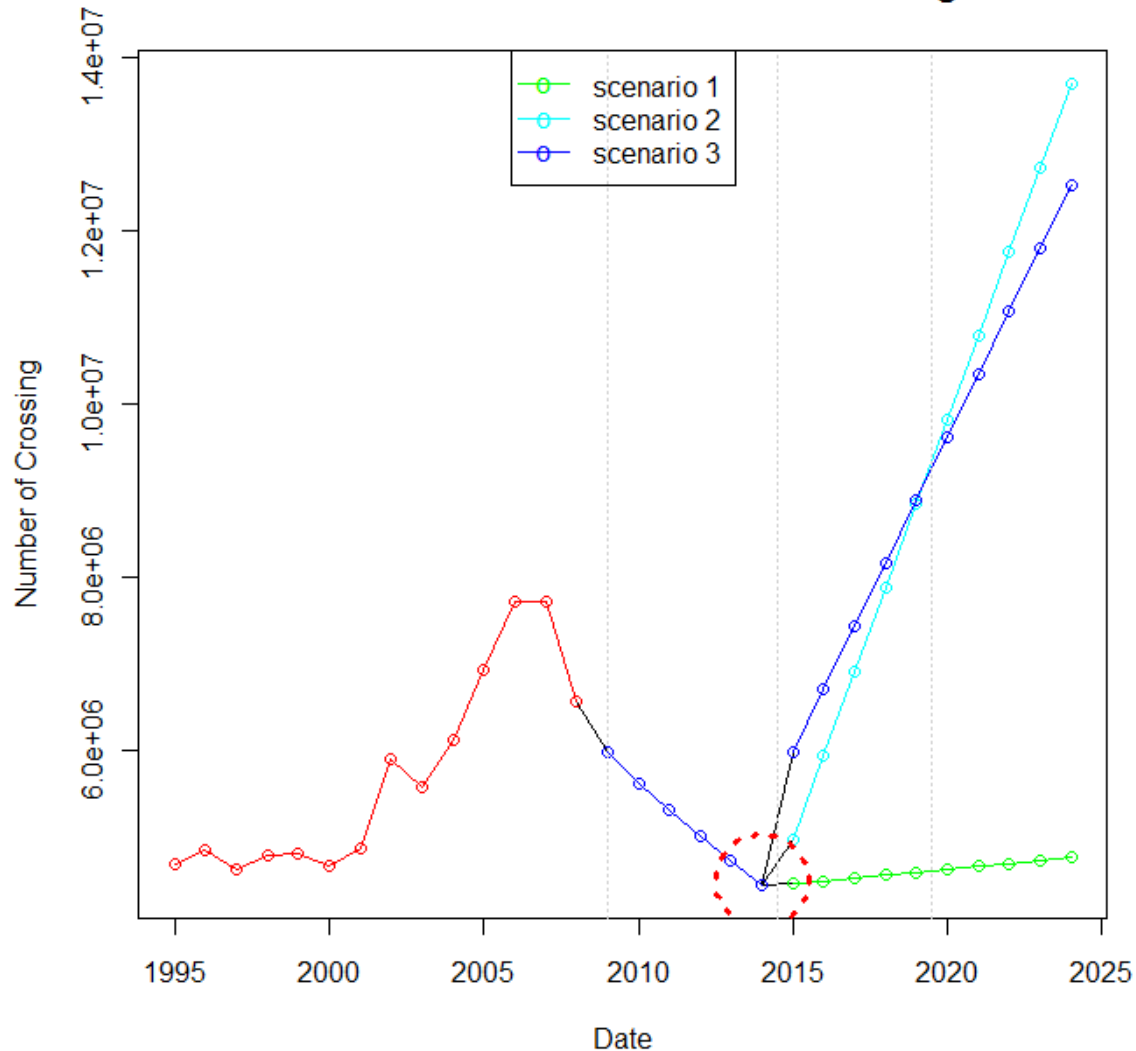
Pedestrian: 5-year Forecast

5-Year Forecast of the Ped Crossing



Pedestrian: 10-year & 15-year Forecast

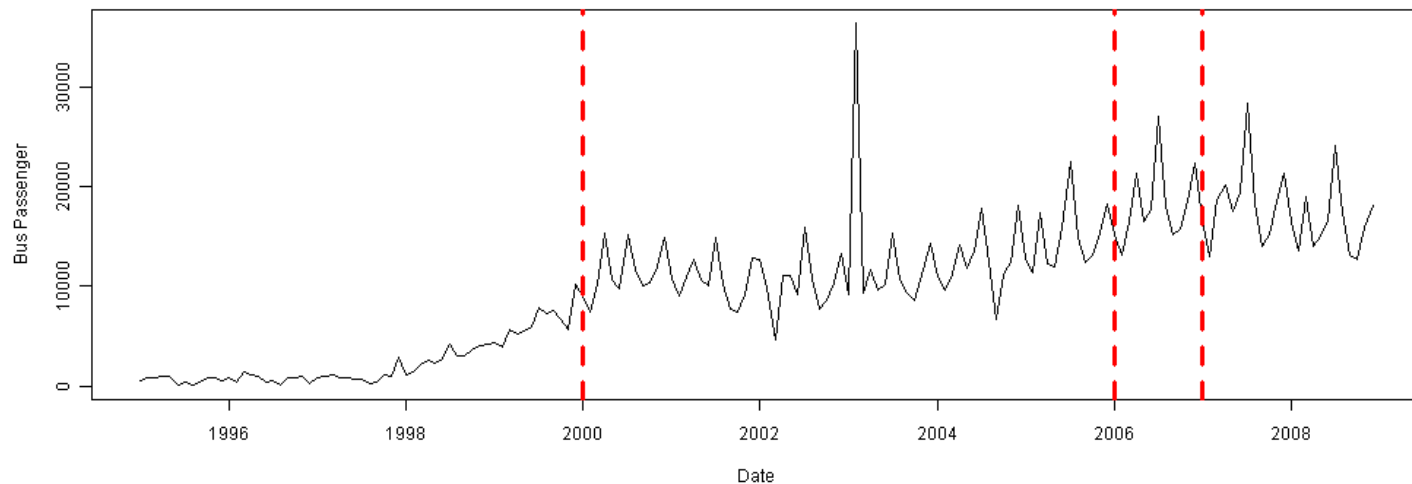
10&15-Year Forecast of the Ped Crossing



Bus Passenger: Forecasts

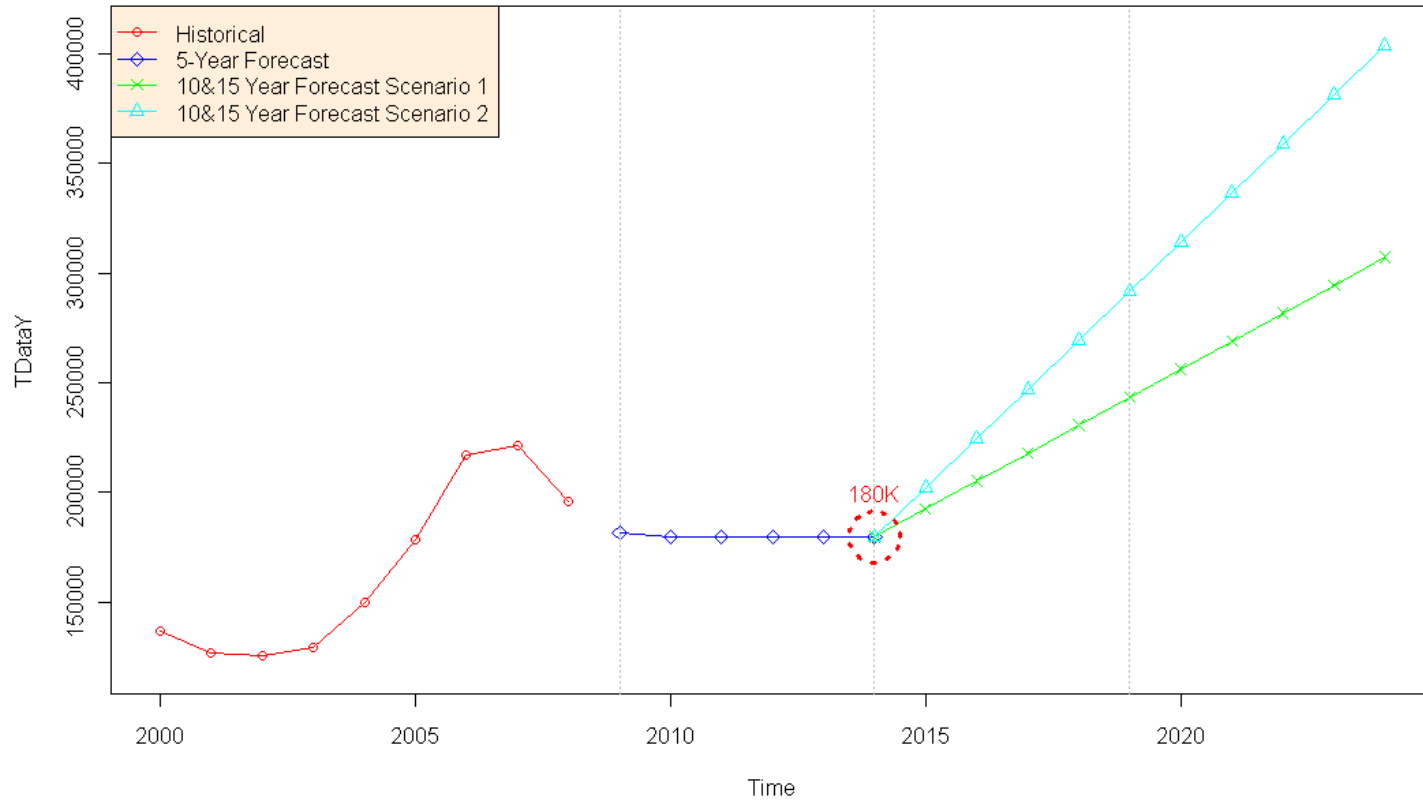


Bus Passengers



- The number of bus passengers after 2000 was very different from that of before 2000
- The number of bus passengers tends to decrease since 2007
- We use data from different time segments to build different scenarios: full data & data between 2002 and 2007

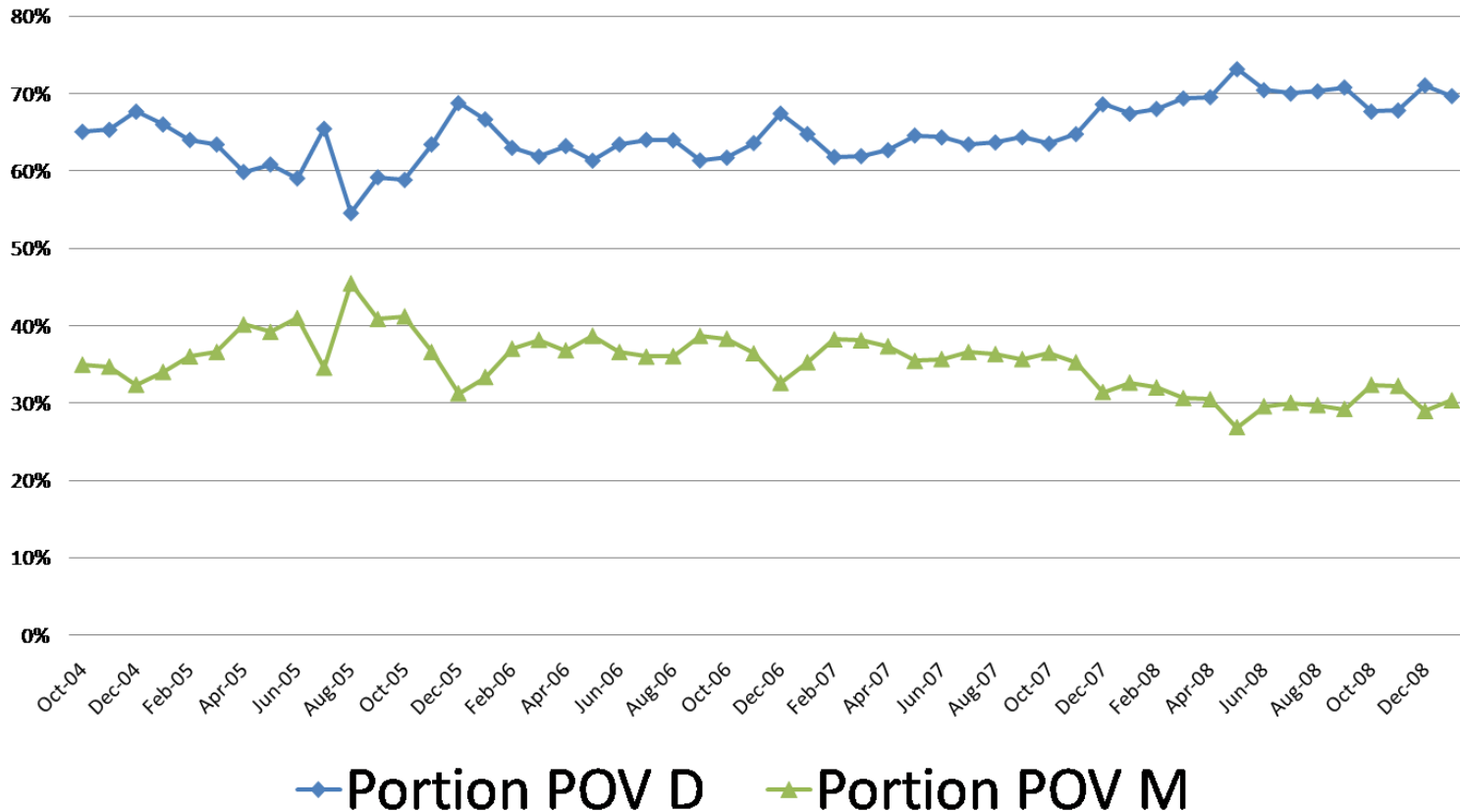
Bus Passengers: Forecast



FORECAST AND DISCUSSION

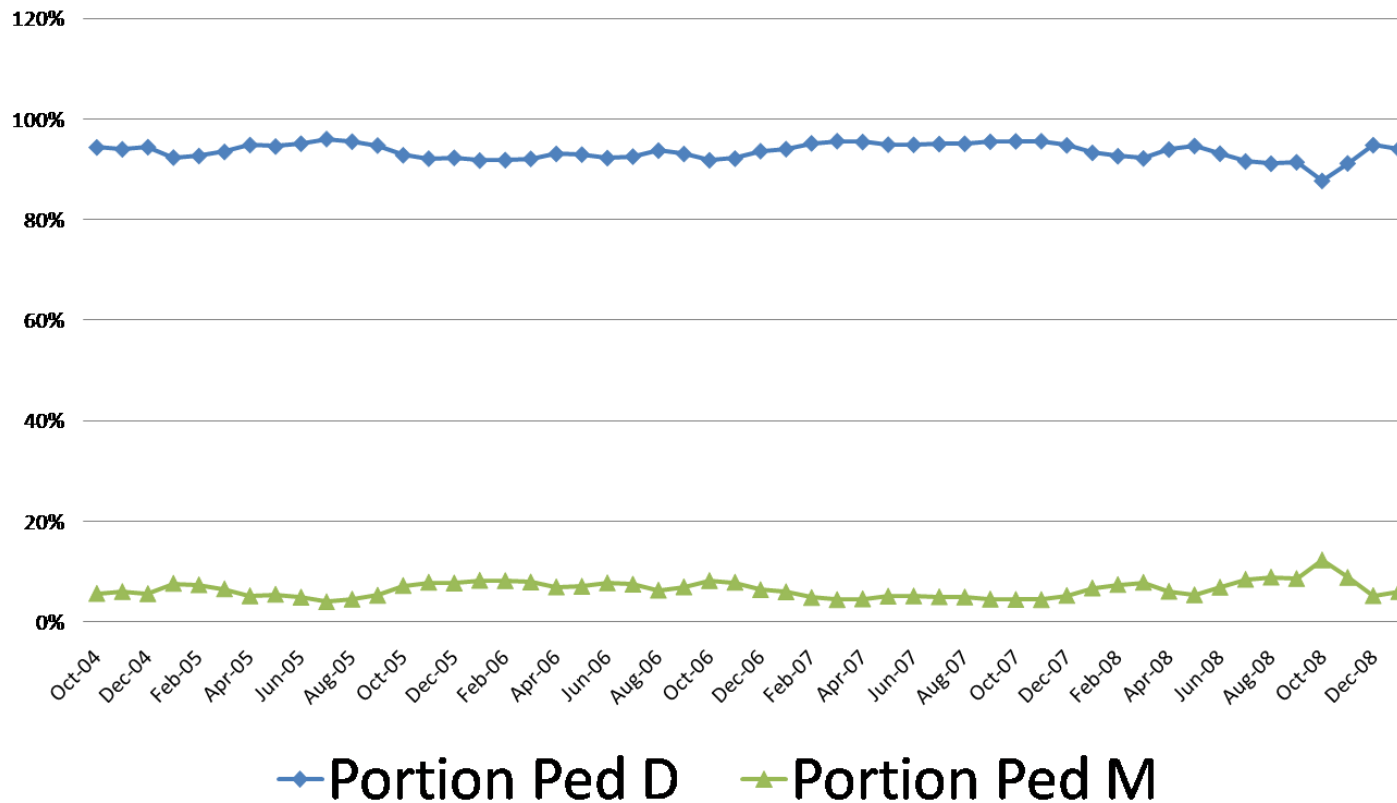
- Traffic Split

POV: Traffic Split



- Before 2007, the portion is roughly 60:40
- Since 2008, the portion is roughly 70:30
- Both of the portions are quite stable

Pedestrians: Traffic Split



- The portion roughly maintain to 95:5

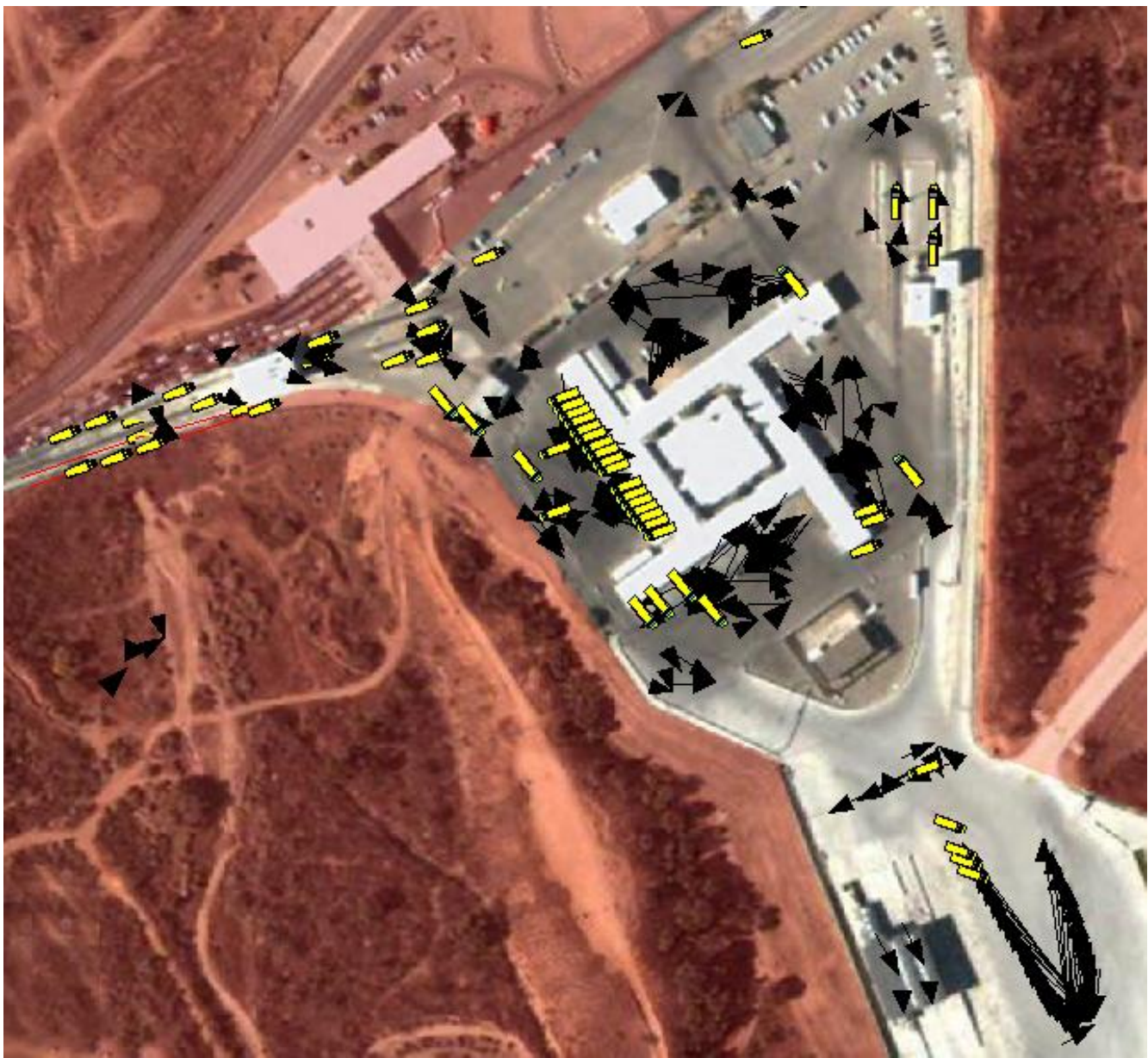
SIMULATION

Overview

- Our model is an updated version of the model used in the ADOT project entitled *Logistics Capacity Study of the Guaymas-Tucson Corridor* (Villalobos et al.)
- Updates were made based on two criteria:
 - Physical infrastructure changes to the Mariposa POE since the previous study
 - Truck crossing times recorded on our visit to the Mariposa POE on May 29, 2009

Infrastructure and Process Changes

- Incorporated 4 lanes and inspection stations throughout primary inspection area
- Designated one highway lane as “fast”, with a potentially different inspection time and direct routing to highway after primary inspection
- Updated assumptions based on field visit
 - Primary inspection times virtually identical between “fast” and regular lanes
 - 30.74% of trucks use “fast” lane
 - Time in CBP facility increased by 7 minutes, which was allocated among CBP inspection areas (details in appendix)
 - Inspection frequencies in appendix



Evaluating Capacity

- Utilized May monthly forecasts with levels for 2014, 2019, and 2024.
- Our evaluation assumes relatively level daily demand throughout the week, consistent with our findings.
- Calculated required processing, average waiting time, and queue length for several scenarios of exchange rate and IIP levels
- Also determined bottleneck locations (primarily Superbooths)

Simulation: Results 15-year Forecast

Scenario	# Trucks	Required Process time	Extra hours required	Avg. time in system (min)	Max in Queue (low 95%)	Max in Queue (high 95%)	Bottleneck	Approx. Utilization
3-1	2302	18.39	7.39	458.475	2262.94	2270.06	SBS	87.69%
3-2	2139	17.21	6.21	426.991	2098.73	2107.67	SBS	81.43%
3-3	2042	16.65	5.65	412.149	2000.89	2008.31	SBS	81.44%
3-4	2325	18.82	7.82	471.270	2285.52	2291.08	SBS	87.71%
3-5	2159	17.28	6.28	433.375	2119.19	2127.61	SBS	83.49%
3-6	2062	16.70	5.70	416.790	2020.59	2030.21	SBS	87.21%

Simulation: Conclusions

- 95% confidence levels on maximum queue length are relatively narrow, and projected maximum lengths vary from 2000 trucks to 2300 trucks
- Results are supported by the observed congestion at Mariposa POE

CONCLUSIONS

Conclusions

- The traffic characteristics at the POEs at Nogales are very different from that of other POEs. One significant difference is the seasonality pattern shown in the truck traffic.
- Our model has face validity, as seen in the predicted vs. real results for the first 7 months of 2009

	Predicted	Real
Jan	29,968	29,667
Feb	29,458	27,926
Mar	30,329	28,952
Apr	27,974	29,773
May	30,104	26,213
Jun	21,819	22,779
Jul	14,935	14,712

Conclusions

- Our forecast is for truck crossings to increase from 30% to 50% over the next 15 years, depending on levels of economic activity and overall movements in exchange rates.
- Vehicles and pedestrian flows are also likely to increase, but these crossings appear to be highly contingent on economic activity levels and are more difficult to specify.

Future Research Topics

- POV traffic has been shrinking since 9/11. Why?
- Does inadequate infrastructure play a part in the shift from vehicle to pedestrian traffic which seems to have occurred in Nogales?
- Improved scenario generation using Delphi techniques
- The usefulness of a central repository for Arizona border studies, projections, and plans.



APPENDIX

Variable Selection: example

Results from exhaustive test (partial)

#	Model	R Square	Test R Sq	VIF	VIF	VIF
78	Truck~USIIP+Xrate	0.9675	0.6710	2.8646	2.8646	
50	Truck~MXIIP+Xrate	0.9671	0.6558	2.2812	2.2812	
19	Truck~AZemp+sonpop	0.9711	0.6524	8.2115	8.2115	
11	Truck~USIIP	0.9667	0.6388			
62	Truck~RXrate+USIIP	0.9668	0.6342	1.3426	1.3426	
4	Truck~MXIIP	0.9667	0.6331			
44	Truck~MXIIP+RXrate	0.9668	0.6279	1.2049	1.2049	
6	Truck~RXrate	0.9668	0.6201			
12	Truck~Xrate	0.9668	0.6043			
25	Truck~AZpop+MXIIP	0.9711	0.5786	3.0764	3.0764	
46	Truck~MXIIP+sonpop	0.9709	0.5681	2.2072	2.2072	
124	Truck~AZemp+sonpop+USDiesel	0.9714	0.5636	8.6778	9.0878	3.5406

- 14 years data & 10 candidate variables. Variable selection is necessary
- Combinations with high value of VIF values were removed
- Two variables in the regression model seems a good choice
- We found that the US IIP and Exchange rate were good variables to incorporate into the model

Strategy of choosing the models

- For Regression Model
 - Enumerate all possible subset, choose the good ones
- For ARIMA model
 - Define a range of each parameters
 - Enumerate all the possible combinations of the parameters within its ranges
 - Generate some relatively good models for future use
- Note that we are not always choosing the best models according to the criteria we defined
 - The criteria may not be able to fully reflect the performance
 - A “too good” performance on training set may lead to over fitting in the forecast
 - Some other issues that not incorporated in the model, but do need to be considered
 - The criteria gave us guidelines, but we cannot only rely on them

Strategy of choosing the models (cont.)

- For each type of model, we selected one of the “best” models of this type, based on statistical criteria
- Compared different type of models
- Chose a type of model to use in the forecast

Updates from simulation – CBP time

- The CBP time measured on our visit of 27.117 minutes was greater than the 20.23 minutes in the original version of the simulation
- To make up the 6.887 minute difference we multiplied the inspection times of each area in CBP by a ratio as calculated below:

Inspection	% of trucks that receive each inspection	calculation	ratios
DOC	83	$(83/133) \times 6.887$	= 4.298
XRAY	33	$(33/133) \times 6.887$	= 1.708
ENFORCE/FULL	17	$(17/133) \times 6.887$	=0.880
TOTAL	133		6.886

Percentage of Trucks Requiring each inspection

Percentage	Description
100 %	Pre-Screening
100 %	Primary Inspection
30.74 %	Released to enter the US from Primary inspection (fast lane)
69.26 %	Required further inspections and enter the compound (normal lanes)
<u>Out of the 62.26% that require more inspection:</u>	
33 %	Required X-Ray
17 %	Required Full Inspection or Hazardous and Weapons Inspection
83 %	Required Documentation Review
20 %	Required to enter the ADOT yard for Inspection