

A
L
A
N
R
O
B
I
N
S
O
N

**VALIDATING TRAFFIC MODELS USING LARGE-SCALE
AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA**

ALAN ROBINSON

A thesis submitted as partial fulfilment of the requirements for the degree of

MASTER OF ENGINEERING (TRANSPORTATION ENGINEERING)

In the

FACULTY OF ENGINEERING

UNIVERSITY OF PRETORIA

November 2017

2
0
1
7

ABSTRACT

Title: VALIDATING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

Author: Alan Robinson

Supervisor: Prof Christo Venter

Department: Civil Engineering

University: University of Pretoria

Degree: Master of Engineering (Transportation Engineering)

Traditional manual survey methods for collecting reliable origin-destination data to develop large strategic transport model is notoriously expensive and the sample sizes are often relatively small. Arguably, the least reliable data required for the development of strategic traffic models is the origin-destination data. Recent technological advances, such as probe data from on-board devices, have been successful in providing data for some needs such as journey times and routing options. However, varying degrees of success have been achieved in obtaining reliable origin-destination (OD) data from these new technologies.

Automatic Number Plate Recognition (ANPR) is one of the newer technologies that could be used to collect large-scale data sets over the large study areas that strategic traffic models cover. The aim of this study is to examine ANPR data collected from the Gauteng Freeway Improvement Project's (GFIP) Open Road Tolling (ORT) gantries in terms of its accuracy and uses in the development and improvement of strategic traffic models. Of particular interest is the use of the ANPR data to contribute towards the improvement of the distribution of trips in the OD matrices. This is achieved by developing methodologies to derive comparable gantry to gantry traffic volumes from the ANPR data and the GFIP traffic model.

The above comparisons enabled the undertaking of a post opening project evaluation of the GFIP traffic model's 2015 forecasts using as many characteristics of the traffic flows and patterns that can be derived from the ANPR data. Characteristics such as traffic volumes and journey times are directly comparable with standard traffic model outputs. Tracking vehicles between gantries enabled the calculation of the number of trips that travel between gantry pairs giving rise to gantry-to-gantry (G2G) trips, which can be represented in a G2G count

matrix. This G2G count matrix has probably the most beneficial data that can be derived from the ANPR systems as it contains an “accurate” element of the trip distribution on the road network.

A methodology was developed to derive equivalent trip matrices from a traffic model’s select-link trip matrices where the links are those where the gantry (ANPR camera) is located. The sums of the trips in the derived sub-matrices match the G2G counts. This enabled the comparison between the modelled trip distribution represented by the select link to select link (SL2SL) volumes and the actual ANPR G2G counts. This is in fact a comparison of a portion of the model’s distribution to actual, comprehensive data.

This study demonstrates that ANPR data has the potential to improve strategic traffic models. The automation of the processes to derive the SL2SL assigned volumes from the models and combining it with existing matrix estimation techniques will enhance the trip distribution in the output trip matrix. The current practice of using individual traffic counts in matrix estimation has the adverse tendency to affect the trip distribution. Hence, the recommendation to use traffic counts in matrix estimation to traffic counts with caution.

LIST OF ABBREVIATIONS

ANPR	Automatic Number Plate Recognition
AVI	Automatic Vehicle Identification
COC	Central Operations Centre
CTO	Comprehensive Traffic Observation
DSRC	Dedicated Short Range Communication
G2G	Gantry to gantry
GPS	Global Positioning System
GSM	Global System for Mobile devices
ID	Identification
kph	Kilometres per hour
OD	Origin-destination
ORT	Open Road Tolling
POPE	Post Opening Project Evaluation
PPP	Public Private Partnership
SANRAL	South African National Roads Agency (SOC) Limited
SL	Select link
SL2SL	Select link to select link
TCH	Transaction Clearing House
TLFD	Trip length frequency distribution
VDF	Volume delay function
VLN	Vehicle licence number
vph	Vehicles per hour

TABLE OF CONTENTS

1	INTRODUCTION	1-1
1.1	Study background	1-2
1.2	Problem statement	1-4
1.3	Study objectives.....	1-5
1.4	Methodology and scope of the study	1-5
2	VALIDATION OF STRATEGIC TRAFFIC MODELS.....	2-1
2.1	Journey Times	2-2
2.2	Traffic Counts.....	2-3
2.3	Trip Distribution.....	2-5
2.4	Summary of Model Validation Methods	2-5
3	LITERATURE REVIEW	3-1
3.1	OD Trip Matrix Estimation	3-1
3.2	New Data Uses in Transport Modelling.....	3-2
3.3	Uses of ANPR (AVI) Data	3-4
3.4	Summary.....	3-6
4	METHODOLOGY FOR USING ANPR DATA FROM ELECTRONIC TOLL COLLECTION.....	4-1
4.1	Source, extent and availability of data.....	4-1
4.2	ANPR Data Extraction	4-3
4.3	Gantry Spacing	4-4
4.4	Processing the ANPR Data.....	4-5
4.4.1	Extracting Traffic Counts.....	4-5
4.4.2	ANPR Data accuracy	4-6
4.4.3	Comparison to other big data sources	4-9
4.5	Gantry to gantry data	4-10
4.5.1	Extraction of trip data	4-10
4.5.2	Gantry to gantry count matrices.....	4-12
4.5.3	Gantry to gantry journey times	4-14
4.5.4	Gantry to gantry speeds.....	4-14

4.6	Summary.....	4-14
5	DEVELOPMENT OF A METHODOLOGY FOR MODEL VALIDATION USING ANPR DATA	5-1
5.1	G2G Formula development	5-1
5.2	Modelled Partial OD to G2G formula	5-4
5.3	Formula testing.....	5-7
5.3.1	Freeway network model.....	5-9
5.3.2	Preparation of Select Link (SL) matrices.....	5-9
5.3.3	Identification of gantry associations	5-9
5.3.4	G2G Matrices.....	5-10
5.3.5	Assignment of SL2SL matrices	5-10
5.4	Alternative formula for complex networks.....	5-12
5.5	ANPR Data for Improved Transport Model Development	5-15
6	CASE STUDY USING THE GFIP TRAFFIC MODEL	6-1
6.1	Traffic Model Description	6-1
6.2	Design Year model scenarios	6-3
6.3	2015 Gantry to Gantry Outputs	6-3
6.3.1	Gantry Counts	6-4
6.3.2	Gantry to Gantry Travel Times	6-4
6.3.3	Gantry to Gantry Travel Speeds.....	6-5
6.3.4	Gantry to Gantry Counts	6-6
6.3.5	Trip length frequency distributions.....	6-7
6.4	Comparison to model outputs and model improvement options	6-7
6.4.1	Traffic Volumes	6-8
6.4.2	Journey times and speeds.....	6-11
6.4.3	Gantry to gantry (G2G) counts	6-13
6.4.4	Trip Length Frequency Distribution	6-15
6.5	Potential GFIP Model Improvements.....	6-16

7	CONCLUSIONS.....	7-1
8	RECOMMENDATIONS FOR FURTHER RESEARCH.....	8-1
9	REFERENCES	9-1

LIST OF TABLES

Table 1: ANPR data entries per month.....	4-3
Table 2: Gantry to gantry distances.....	4-4
Table 3: Example of Counts from Gantries.....	4-5
Table 4: ANPR Data Errors.....	4-9
Table 5: Example of the Association Table of Up- and Down Stream Gantries.....	5-10
Table 6: G2G Combinations from Routes.....	5-14
Table 7: February 2015 Average Weekday Counts at Gantry 6, 7, 8 and 9.....	6-4
Table 8: February 2015 Weekday AM Peak Hour G2G Travel Times (minutes).....	6-5
Table 9: February 2015 Average weekday speeds between 07:00 and 08:00 (kph).....	6-6
Table 10: Gantry to Gantry Counts for All Vehicles on Weekdays - 07:00 to 08:00.....	6-6
Table 11: Comparison of Gantry and Model Traffic Volumes (Average 06:00 to 09:00).....	6-9
Table 12: Validation Criteria Results.....	6-11
Table 13: Summary Statistics for G2G and Model Count Comparison.....	6-14
Table 14: Measured and Modelled Average Trip Length on the GFIP Network.....	6-15
Table 15: Comparison of percentage of trips versus distance travelled.....	6-16
Table 16: Worst GEH Results from G2G Comparisons.....	6-19

LIST OF FIGURES

Figure 1: GFIP Freeways and Gantry Location (Source SANRAL).....	4-2
Figure 2: Comparison of ANRP (Gantry 8) and CTO Data (Station 1863).....	4-7
Figure 3: Example of the average G2G counts for all vehicles on weekdays between 07:00 and 08:00.....	4-13
Figure 4: Example of an average G2G travel time matrix on weekdays between 07:00 and 08:00.....	4-14
Figure 5: Select Link Trips Passing Gantry B.....	5-2

Figure 6: Modelled Freeways Assigned AM Peak Hour 2015.....	5-9
Figure 7: Trips between Gantry 13 and 7.....	5-11
Figure 8: Trips between Gantry 8 and 23.....	5-11
Figure 9: Select Link at Gantry 23.....	5-12
Figure 10: Trips between Gantry 8 and 23 only.....	5-12
Figure 11: G2G Trip Length Distribution.....	6-7
Figure 12: G2G Trip Distribution without Diagonals.....	6-7
Figure 13: Comparison of traffic volumes by vehicle type and location.....	6-10
Figure 14: Model versus Gantry Journey Times on the N1.....	6-12
Figure 15: G2G vs Model Output - Light Vehicles.....	6-13
Figure 16: G2G vs Model Output - Heavy Vehicles.....	6-14
Figure 17: ANPR and Model Trip Length Frequency.....	6-16
Figure 18: Select Link Matrix ODs through Gantry 13.....	6-20
Figure 19: G2G Matrices from the GFIP Model and Actual Values.....	6-21

LIST OF MATRICES

Matrix 1: Select Link from location A.....	5-2
Matrix 2: Select Link from location B.....	5-3
Matrix 3: Select Link from location C.....	5-3
Matrix 4: Select Link from location A, B and C.....	5-3
Matrix 5: Select Link from location A and B.....	5-4
Matrix 6: Step 1 result with cells AB and ABC.....	5-5
Matrix 7: Step 1 less upstream SL_{B+1}	5-5
Matrix 8: Resultant Matrix of Trips through Only Location A and B.....	5-6
Matrix 9: Trips through gantry B including those through A and C.....	5-6
Matrix 10: Removal of trips through gantry A and gantry B.....	5-6

LIST OF APPENDICES

Appendix A: ANPR Data

Appendix B: Gantry to Gantry Distances

Appendix C: Data Miner Software

Appendix D: Preparation of Gantry Data

Appendix E: ANPR Gantries and Electronic (CTO) Count Stations

Appendix F: Comparison of ANPR and CTO Data

Appendix G: Modelling Process for Formula Validation

Appendix H: Gantry / Model Comparisons

ACKNOWLEDGEMENTS

I would like to thank the following for their support and assistance in the preparation of this Thesis.

- Professor Christo Venter for his guidance and encouragement
- SANRAL for the data upon which this work is based and the permission to use the traffic model developed for the GFIP traffic and toll revenue forecasting
- My wife and children for their continued support and encouragement throughout the preparation of this work.

DISCLAIMER

The views, recommendations and inferences contained in this thesis are those of the author and not necessarily those of the University of Pretoria or the South African National Roads Agency (SANRAL).

1 INTRODUCTION

A significant amount of investment is required for transport infrastructure such as high mobility roads, high-speed trains or dedicated bus ways. Where funding is not readily available from the fiscus, financial and economic assessments are used to determine the ability to fund the projects using alternative funding mechanisms. Governments allocate funds based on the economic analysis, but when additional funding is required, governments can resort to Public Private Partnerships (PPPs) and raising funds on the open market. Funding agencies lend money to project sponsors against the project's potential to deliver an independent, reliable and long-term revenue stream. The revenue streams are estimated from patronage/traffic and fare/toll revenue forecasts over the project's life, derived from traffic (multimodal transport) models developed for the project. The quantum of funding that lenders will invest into a project depends on two factors, the value of the revenue stream and the confidence the lenders have in the models used to derive the revenue streams.

For toll road projects, accurate traffic forecasts are required to calculate the project's toll revenue stream as well as the project's initial and future infrastructure requirements, which play a major role in calculating the project's cost stream. Therefore, traffic forecasts are critical inputs into the project's financial model from both the cost and revenue perspectives. It is therefore understandable why so much emphasis is placed on the accuracy and robustness of the traffic models. Model accuracy reduces the potential risk related to producing traffic/passenger forecasts to calculate the economic benefits and/or the cost and revenue streams that are input to financial models. This is important to ensure that they are "investment grade" and minimise the risk to the lenders and project sponsors (Bain, 2009).

A fundamental input into a traffic model is reliable origin-destination (OD) data, yet this data is arguably the least comprehensive data used in the development of a traffic model. It is typically expensive to obtain and traditional survey practices produce relatively small samples. The OD data forms the primary input into the model's trip matrices, which define the number of trips that travel between origins and destinations, or traffic zones (spatially distributed residential and activity areas in a city).

Large traffic models can contain thousands of traffic zones and the number of matrix cells, each representing one OD pair, is the square of the number of zones. Obtaining enough data to fill each cell is practically impossible. Sample OD data is analysed to derive a distribution function that mathematically estimates the trips in each cell based on the generalised cost of travel between zones and the numbers of trip generations from, and attractions to, them.

The model calibration process entails adjusting the various parameters and the trip matrices, such that the traffic patterns produced by the model represent the actual traffic conditions. Adjusting the model to one data set can negatively affect other elements of the model, for example, adjusting the trip matrices based on traffic counts can affect the matrix trip distribution, thus the trip length frequency distribution, and vice versa. The validation of the model compares the calibrated model outputs to independent data sets not used in the model calibration. Typically, a percentage of independent traffic counts are excluded from the calibration for this purpose.

New technologies have resulted in significant quantities of data that is available for use in transport engineering, planning and modelling. There is however, a tendency amongst engineers to develop algorithms that are more complex purely because this data is available. Ermentrout (2002) asks the question “At what point does the model cease to have explanatory value?” Often, the more complex the model, the more input assumptions are required, which can make it difficult to test the sensitivity of inter-related assumptions. Wu (2013) suggests that using the additional data with existing relatively simple algorithms could produce a better solution.

The Gauteng Freeway Improvement Project is a toll road project that utilises ANPR technologies for the identification of vehicles passing under the toll gantries deployed along the upgraded freeway network in an Open Road Tolling (ORT) system. Large quantities of ANPR data was made available for analysis and to determine if it could be used in transport modelling.

The overall aim of this work is to investigate the use of large data sets obtained from automatic number plate recognition (ANPR) systems to improve the robustness of, and confidence in, traffic forecasting models using current transport modelling software platforms.

1.1 STUDY BACKGROUND

The traditional four-step traffic/transport model is a mathematical tool that combines all trips in a study area (Ortúzar & Willumsen, 1998). The purpose of the traffic model is, firstly to represent the current travel patterns and then to predict outcomes resulting from various stimuli in the transport system, ranging from population growth, road network expansion to the introduction of public transport services. The four steps in the model development are:

- Trip generation – defines the quantum of trips that originate from or are attracted to an area (traffic zone). The calculation uses land use extents and types with trip generation rates, both data sets are readily available.
- Trip distribution – defines the function to derive the propensity of trips between zones. These distribution functions are derived from origin-destination (OD) surveys.

- Modal split – calculated to proportion of trips that are transported by the various modes of transport, i.e. cars, bus, taxi, train, walk, cycle.
- Trip assignment – the process of calculating the quantum of trips on the various elements of the transport network, i.e. sections of road and on the public transport system.

Obtaining sufficient reliable OD information is a challenging exercise in transport modelling. Traditional survey techniques (van Vuren, Clarke, & Davidson, 2004) for obtaining limited amounts of OD information include:

- Household interviews
- Travel diaries
- Road-side interviews
- Manual number plate surveys

Each of these survey methods are cumbersome, expensive, susceptible to sampling bias and human capture errors, and limited in coverage.

Technological advances have presented practitioners with the means to collect trip data electronically. This includes the ability to track individual vehicles in real time, giving rise to the term “big data” (in traffic and transport modelling terms). These techniques include:

- Global System for Mobile (GSM) tracking
- Global Positioning System (GPS) tracking
- Bluetooth tracking
- Automatic Number Plate Recognition capture (ANPR)

The GSM and GPS tracking are probe-data technologies, i.e. the monitoring of the movement of an in-vehicle device. Bluetooth tracking and ANPR surveys monitor targeted entities that pass specific locations and track them from one point to the other using either an in-vehicle Bluetooth device or the vehicle’s number plate. ANPR data can capture every number plate from passing vehicles and is therefore a very accurate and efficient number plate survey.

The opportunity for this research followed from the completion of the Gauteng Freeway Improvement Project (GFIP); the upgrading of the province’s urban freeways to address the urgent need to rehabilitate the road pavements and to add capacity to the road network. The South African National Roads Agency (SOC) Limited (SANRAL) implemented this project between 2006 and 2011. The original intention was to fund the project through the collection of tolls using open road tolling (ORT) equipment. This equipment included 42 one-way toll

gantries located along the 201km of upgraded freeway at approximately 10km intervals. When a vehicle passes under a gantry, the system records the following information:

- Information from the e-tag (if present) that includes the number plate, vehicle class and fund balance
- Photographs of the front, back and overall position of the vehicle
- The vehicle number plate using ANPR technology
- The exact time stamp and
- The vehicle's toll classification based on the volumetric profiling of the vehicle.

Traditionally, a large manual number plate survey would result in samples of between 15% and 90% depending on the volume of traffic travelling on the surveyed road. Adding in recording and capture errors, the matching pairs can be as low as 5% of the sample. The primary function of the ORT system is to collect tolls from all vehicles travelling under the gantries. Since the accuracy of the financial transactions depend on information collected through via the toll gantries, a high degree of accuracy is a requirement in the deployment and operations of the ORT system. In comparison to traditional survey methods, this data is would be significantly large and incorporates traffic volumes at the gantry locations, travel times (and speeds) between gantries and the number of trips that travel between specific gantry pairs giving a spatial representation of trips through the freeway network.

The ability to improve the confidence of the traffic and toll revenue forecasts by using very reliable data will improve investor confidence in transport schemes such as toll roads. Incorporating comprehensive and reliable data into the development, calibration and validation of the traffic models should assist in the confidence of the results derived from these models. The GFIP traffic model was develop before the ANPR data became available and the ANPR data this study uses to validate the model forecasts.

1.2 PROBLEM STATEMENT

The predominant focus in the use of big data sets is in real time traffic and incident management systems and driver information. Its use in the development of strategic traffic forecasting models has had varying degrees of success. For example, probe data are samples of moving devices and whilst it is possible to track the probe vehicles between their assumed origins to destinations, converting these samples into person or vehicle trips to represent the vehicle population during the modelled time-periods require expansion factors, which in themselves are a potential source of error. Bluetooth tracking is a similar concept to ANPR data where Bluetooth devices are

recorded passing receiver stations along a route or throughout a network. The problem of relating a device to a person, vehicle or numbers of people in a bus remains a potential source of error.

ANPR data in this study, on the other hand, entails the recording of all vehicles at the recording sites over time. The problem remains how to incorporate this data into the development of robust transport models effectively. Whilst numerous papers describe the use of ANPR data to calculate journey times for use in model network calibration and validation, the ability to relate this data to modelled OD data is limited to where the ANPR stations are on a closed cordon such as on freeway entry and exit points or on inter-city routes.

1.3 STUDY OBJECTIVES

This study addresses the problem of using ANPR (or Bluetooth) recording stations in an open layout where there are multiple entry/exit points and zones between stations. The problem is to isolate the trips within an OD trip matrix that route through various ANPR stations whereby they are first recorded at one station and last recorded at another station before reaching their destination.

The objectives of this study are therefore to:

- Examine the extent and characteristics of the data obtained from ANPR systems using the GFIP ORT data as a case study
- Develop a methodology to derive counts of trips between specific gantry pairs using the ANPR data and compare these to equivalent traffic volumes derived from modelled OD trip matrices.
- Test the above methodology on a case study traffic model to validate the 2015 design-year traffic forecasts, produced from the GFIP 2006 base-year traffic model, against actual data.

A key element of this work is therefore the development of methodologies and processes required to produce comparative traffic information from the ANPR data and the traffic model. The comparison between the two data sets is effectively the validation of the model's output.

1.4 METHODOLOGY AND SCOPE OF THE STUDY

An outline summary of the methodology followed in this study comprises the following elements:

- Describe the requirements of the standard traffic model and the parameters that could be improved using ANPR data (Chapter 2)

- Review current uses of ANPR data in transport modelling through available literature and how the has been used in transport model development (Chapter 3)
- Review available ANPR data, assess the option for processing the data and compare its accuracy to other data sources (Chapter 4)
- Develop the methodology to produce equivalent gantry to gantry traffic volumes from traffic models for comparison to the ANPR data (Chapter 5)
- Test the methodology on a case study model, being the Gauteng Freeway Improvement Project Model developed using 2006 data, comparing the 2015 freeway traffic, operations and distribution forecasts to the ANPR data and determining ways to improve the model (Chapter 6)
- Provide conclusions on the study work and recommend further work to expand on the use of big data in the development of traffic and transport forecasts. (Chapter 7).
- Assess where this work can be extended to enable the calibration of models using ANPR data and recommend further work in this field to maximise the use of this relatively new and underutilised data (Chapter 8).

Whilst the process enables the identification of the trips within the OD trip matrix that corresponds to a traffic count between two specific ANPR recording stations, this process could be extended to include model calibration processes. However, this is beyond the scope of this work.

The process of validating a model using the ANPR data is demonstrated using a 2015 peak hour forecast from a 2006 base year model. Therefore, the validation checks were undertaken to highlight potential GFIP model improvements that can be made in the future updating and re-calibration of the model (not part of this scope). The processes can then be employed to validate the revised base year models including all time periods and vehicle classes.

2 VALIDATION OF STRATEGIC TRAFFIC MODELS

Traffic models are mathematical tools developed to answer specific traffic planning and/or design questions. A model's data requirements depend on the type of model, which can range from wide area strategic planning to detailed intersection analysis, and the detail in the required output. The definition of the trip movements through the model can be represented in the form of a trip matrix. A single length of road will have a 2x2 matrix, each end being the start and end-point (zone) of the trips using the road. A 4-leg intersection will have a 4x4 matrix; excluding the diagonal cells, the 12 remaining cells will contain the trips that travel into the intersection on one leg and out of another. Larger strategic models may have thousands of traffic zones, each cell containing the trips that travel from one zone (represented by the row number) to another zone (represented by the column number).

A 2-way traffic count will provide the contents of a 2x2-trip matrix. Using an electronic traffic counter over an extended duration will significantly improve the accuracy of the average number of trips occurring during the modelled time period. Manual turning movement counts will provide the contents of a 4x4-trip matrix. It is not as straight forward to do these counts over an extended time period and therefore often rely on a 12-hour count undertaken on one or two days. The daily variation of such counts can be significant, which would have an impact on the reliability of the outcome of an analysis.

Obtaining sufficient trip data for the origin-destination matrices of a large strategic model such as the GFIP model, containing approximately 900 traffic zones with over 800 000 cells is considerably more difficult. As mentioned previously, the Gauteng Household Travel Survey (Gauteng Province, 2014) undertook 30 000 interviews of which approximately 15% were car related trips, i.e. 4 500 entries for an 800 000 cell matrix. Breaking this down into user classes dilutes the sample even further. Freight trips were not included in the survey. In conclusion, the base data for an origin-destination matrix is somewhat lacking, hence the interest in electronic methods to collect OD data.

The calibration of a model entails the adjustment of various parameters so that the resultant output is representative of the data used to develop the model. The validation of the model is the comparison between the model output and independent data. There are generally three ways to validate a strategic model:

- Against journey times
- Against traffic counts, and
- Against the trip distribution

This section reviews these validation areas and describes the influence that one process has on the other.

2.1 JOURNEY TIMES

The strategic model network is generally link based. Therefore, the volume delay function for each link must be calibrated to account for the effect of congestion on the link and delays incurred at intersections. Where the model has simulated intersections, the volume delay function excludes junction delays. A standard equation to calculate delay from varying traffic volumes takes the form (van Vliet, 2015):

$$t = t_0 + aV^n \quad (V < C)$$

Equation 1

where:

t = the link time

t₀ = the link time under free-flow conditions (= free-flow speed / distance)

V = the link volumes

C = the link capacity

Constants a and n calculated from the capacity of the link and the average speed that is attained at capacity. The volume –delay function input required by the SATURN model includes free-flow speed, speed when the volume reaches capacity, the link capacity and the constant “n”, which defines the shape of the curve. Higher values of n mean that the free-flow conditions remain until just before reaching capacity as found on freeways. Lower values result in speeds beginning to reduce at volumes much lower than the link capacity.

The calibration of the volume delay functions uses data collected from specific locations where speed and volumes are collected for a link type. However, it is assumed that the derived volume delay function is constant over all similar links. The function must apply to all time periods and account for normal road gradients. Factors such as slight gradients, sight distances and intersection type and spacing result in numerous variations. The aim of the calibration of the volume delay functions is to adjust the function parameters and constants to best reflect the measured journey times on specific road (link) types in the traffic model.

The validation of the modelled network is the ability to demonstrate that the modelled journey times correlate to the measured journey times along the key routes under prevailing traffic

conditions. The total journey times are the cumulative journey times along the links on a route; it is a measure indicating that the modelled network is functioning in an acceptable manner. The correct functioning of the network is important as it promotes relisting routing of traffic through the network, an extremely important factor when developing trip matrices that match observed traffic counts.

A validated network does not mean a calibrated model. It is possible to calibrate and validate a network starting with a unit matrix, i.e. no trip distribution function in the initial trip matrix.

2.2 TRAFFIC COUNTS

Calibrating and validating a traffic model to traffic counts is a common occurrence and unfortunately necessitated because of insufficient origin-destination data. Van Vliet (2015, pp. 13.1 - 13.6) provides a methodology that can synthesise a trip matrix that, when assigned to the network will reproduce observed traffic flows on a link. The trip matrix T_{ij} contains all trips that travel between origin i and destination j and the sum of the proportions of these trips that use a link a would result in the observed link volume, i.e.:

$$\sum_{ij} T_{ij} \cdot P_{ija} = V_a^{obs}$$

Equation 2

where:

T_{ij} = the output matrix

P_{ij} = the proportion of trips from i to j passing along link a

V_a^{obs} = the observed volume travelling on link a

The process incorporates measures to ensure the result is as close to an original trip matrix as possible using an entropy maximising model to seek a solution. An iterative process finds a set of balancing factors X_a for link “a” with a traffic count such that the assignment of the overall matrix matches the observed traffic count. In the SATURN matrix estimation software, additional matrix constraints are possible including “fixing” specific cells, fixing certain trip ends and/or limiting the amount of change permissible for row and column (trip end) totals. The process is iterative and similar to the Furness matrix balancing process and processes each counted link in turn. The processing of one counted link can be “unbalanced” by the processing of a second link that includes trips that pass through a previously counted and processed link.

Using this matrix estimation process introduces common modelling errors when the following is evident:

- Not having a calibrated network. If the routing of trips is incorrect, the wrong OD pairs may pass through the counted link and, when factored, cause distribution errors in the matrix.
- Using unbalanced counts. When trips of an OD pair pass through one counted link with a count higher than the assigned volume, an increasing factor X_a is applied. In the next link calculation, a count lower than the assigned flow will result in a factor X_a that will reduce the OD pair. Through the iterative process, the X_a should converge. If the counts are correct and balanced, changing the factors of the OD pairs that only pass through one of the counted links will take up the difference. If the counts are not balanced, the result will be unpredictable.
- If the initial OD matrix produces X_a factors that are all either high or low, the matrix estimation process can distort the trip distribution by only factoring up or down those parts of the matrix that have trips pass through counted links.

Whilst using traffic counts to calibrate a strategic traffic model is a common and well documented process, all it results in is the knowledge that the model predicts the information provided to it. Validating the model to fully independent traffic counts provides some confidence that the model is reflecting the correct volumes in parts of the model that did not rely on calibration data. Whilst, this is a good indication that the route choices through the network is functioning in a realistic manner and that the matrix has a representative number of modelled trips, it does not necessarily relate back to the trip distribution.

Using traffic counts for model calibration and validation can present modellers with a dilemma. When obtaining traffic counts an objective is to obtain as many counts as possible. The separation of the counts into those for the calibration and those for the validation of the model is the first problem. Common practice is to keep 15% of the counts for validation, however, when one has a very good coverage of counts, which 15% are “independent”? With a good coverage, most counts are influenced by other up- or downstream counts. As a result, using too many counts for the model calibration compromises the validation of the model’s trip distribution. It will however provide a very good validation of the trip generation rates used for converting land use information in large and potentially mixed-use zones.

2.3 TRIP DISTRIBUTION

A doubly constrained gravity model is commonly used in the development of strategic transport models (van Vliet, 2015), taking the form:

$$T_{ij} = A_i B_j O_i D_j e^{-\beta C} \quad \text{Equation 3}$$

where A_i and B_j are row and column balancing constraints to satisfy the row and column totals O_i and D_j . C is the generalised cost of the trip between O_i and D_j . β is a constant calibrated from survey data.

The calibration of β is critical to the gravity model as it is a factor in the calculation of the average trip length, and this in turn, influences the trip distribution. Calibrating β from survey data is very dependent on the completeness and representativeness of the OD survey data. As mentioned previously, obtaining sufficient representative trip data to calibrate β to any degree of accuracy is problematic and the reason for much of the current research into using electronic means to collect this data.

With the knowledge that various input parameters and trip matrix estimation techniques can alter the trip distribution derived from survey data, the trip length frequency distribution from the calibrate model is compared to that obtained from the limited survey data. If the output of the model calibration processes has remained within acceptable parameters, the mode is deemed validated.

2.4 SUMMARY OF MODEL VALIDATION METHODS

Model validation is the process of determining if a traffic model is representative of the traffic patterns prevailing in the modelled road network. This includes the routing of trips through the network, the number of trips on the road links and the distribution of trips through the network. The calibration and validation of the road network and the traffic counts could be “achieved” without initial trip distribution information if there is a high concentration of traffic counts throughout the modelled study area. This does not mean a validated model.

Calibrating the gravity model from sparse or synthesised survey data to determine an initial trip distribution and then validating the post-calibration trip length frequency distribution to this data is an important step for model validation. The representative distribution of trips in a model is essential, considering the primary use of the model. Without confidence in the traffic

distribution, there is no confidence in the model's ability to predict the traffic volumes that would use a new road scheme, or pass through tolling points. Using the model to determine the patronage of a high-speed rail facility, where the target market is the car user, there would be little confidence in the viability of the project.

The importance of validating a traffic model to comprehensive, reliable and independent trip distribution data cannot be understated. In section 4 above, the accuracy and coverage of ANPR data is demonstrated, however there is still a disconnect between producing a gantry-to-gantry (G2G) count matrix, which inherently contains distribution data, and comparing it to model output in an efficient way. Section 5 describes the development of a method to extract an equivalent G2G assigned count matrix from the calibrated model enabling the model to be validated against comprehensive and reliable distribution related data.

3 LITERATURE REVIEW

Whilst the aim of this work is to determine a methodology to incorporate ANPR data into the development and accuracy of strategic traffic/transport models, specifically emphasis is given to improving the estimation of the trip matrices. To achieve this, one must first understand the fundamentals of estimating trip matrices from available data.

- The development of OD trip matrices from roadside and other survey data
- Using new technology data including ANPR data in traffic models.

This section reviews previous work undertaken in the above regard to establish where additional research is required to further the use of new data in the development of improved transport models.

3.1 OD TRIP MATRIX ESTIMATION

Origin destination (OD) trip matrices are fundamental inputs into traffic studies and transport models. The development or synthesis of trip matrices has been the focus of many studies dating back to the 1970s from John Wootton in 1972 (Kirby, 1979) where observed data only provides information to form partial matrices. Such data was obtained through household, roadside or other survey techniques. The full matrix is derived using distribution functions calibrated from observed data. It was noted at that time that some practitioners were using the partial matrix techniques to “complete” matrices without fully understanding the assumptions required when using them, to the point where their use was inappropriate.

Willumsen (1981), Fisk (1989) and Tamin & Willumsen (1989) consider various methods of matrix estimation from traffic counts to produce lower cost trip matrices. The gravity model, opportunity model and gravity-opportunity models were evaluated in terms of identifying the ODs associated with a link count under congested conditions resulting in multi-routing. These models incorporate techniques including entropy maximisation and generalised least squares estimators (Cascetta, 1984) (Hazelton & Gordon, 2002) and a non-linear programming approach developing code for inclusion in software (Doblas & Benitez, 2005).

An important outcome of this work is that the level of accuracy in the matrix depends on:

- The trip making algorithm in the model used, i.e. all-or-nothing, equilibrium or stochastic assignments
- The method used for matrix estimation
- The reliability and independence of the traffic counts

- The level of detail in the zone system represented in the model
- The accuracy of any prior trip matrix used in matrix estimation methodology.

Ortúzar & Willumsen (1998) show that traffic on a section of road (link) results from the combination of a trip matrix and route choice. Therefore, the sum of all OD pairs in the matrix which use the link must equal the traffic count on the link. A select link analysis will provide the proportion of the trips from each OD pair that passes through the counted link. To create a unique matrix with N zones one would require N^2 fully balanced traffic counts (all taken at the same time with no other sinks and sources other than the zone connectors). This is an impossible task in large scale models. The task is also complicated in congested networks where there are multiple route choices between origins and destinations. However, using a gravity (impedance) model or a prior matrix derived from a previous study, together with any zonal trip-end constraints as a starting point, an acceptable solution can be derived by using entropy maximising techniques. The accuracy of the output matrix is based on the following issues that increase the number of possible solutions:

- The higher the volume of the traffic counts, as on freeway links, the higher the number of OD pairs involved.
- The reduction of the number of counts increasing the number of “un-checked” OD pairs

The objective of acquiring data to optimise the matrix estimation process would be to ensure that:

- The data is reliable and consistent
- Each element of the data is independent of other data
- The number of measurements are maximised
- Each element of data relates to as few OD pairs as possible.

3.2 NEW DATA USES IN TRANSPORT MODELLING

The cost of undertaking traffic surveys is leading to studies to optimise the number and location of traffic counts (Castillo, Menéndez, & Sánchez-Cambronero, 2008). This work was extended to look at more recent technologies relating to Automatic Vehicle Identification (AVI) and in particular ANPR. Minguez et al (2010) and Hadavi & Shafahi (2016) make a case for optimising the location of AVI sensors, which are critical to estimating OD matrices. They consider various models for the selection of location of survey sites based on budgets while prioritising uniqueness of path and/or the maximising paths and OD flows through the observation sites. This is useful when considering the expansion of an ANPR network through a modelled network.

Putting the cost of acquiring survey data into perspective, roadside interviews cost up to €10 per trip record and capture approximately 10% of the passing traffic on busy roads (van Vuren & Carey, 2011) (Pollard, Taylor, & van Vuren, 2013). The average cost per household for household travel surveys ranged between \$100 and \$250 in 2007, with the average being \$131.89 per interview to achieve a sample of 1 000 for a population of 1.5million (Hartgen & San Jose, 2009). The Department for Transport (2014) (DfT) recommends that primary traffic data should not be more than six years old. Therefore, authorities will incur significant costs to maintain sufficient current data for the modelling of a province such as Gauteng with approximately 13.4 million inhabitants in 4.5 million households in 2016 (Statistics South Africa, 2016).

The cost of traditional manual surveys has encouraged researchers and transport planners to review current technology to obtain the required data for transport planning. A summary of research undertaken for Florida, USA (Cambridge Systematics Inc, 2012) covers various technologies including probe measure and spot measure technologies. Whilst most of the report centred on travel time measurements, a summary of the findings with respect to origin-destination data collection includes:

- Toll Tags – only available on toll roads and will require agreements with toll road operators to widen the area of the survey using permanent or mobile tag readers.
- Bluetooth Device Matching – recording the Median Access Control (MAC) address of devices passing detectors. A 5% matching sample rate limits the effectiveness of this technology, especially when used for long distance trips. The location of the readers does not provide the actual origin or destination of the trip.
- Wireless Location Technology - tracks cell phones using the towers and the transmittal of frequencies between towers. The information is obtainable from the cell phone vendors. Because the location tracing is done during calls and location data is not recorded on a continuous basis, and the data can be bias if a company has targeted clientele. This data must be combined with additional land use and demographic data, monitored over a long period to identify repetitive patterns for the data to be useful OD data.
- Crowd Sourcing – derives its real-time data from devices/applications used in the traffic stream. Google Maps (Google, 2009). Other service providers include TomTom© and Waze©.
- Automatic Licence-Plate Capture (ALC) (ANPR) – is similar to Bluetooth but identifies vehicles at sensor locations. Theoretically, ANPR data enables the access to owner information through the vehicle licencing department. This does have privacy of information implications.

There are two general uses of the various technologies; operational trip information related to travel times, and the derivation of OD trip patterns. With the access to ANPR data, this work focuses on the latter and seeks to determine research that has been undertaken in this regard.

3.3 USES OF ANPR (AVI) DATA

In a simple form, the concept of using ANPR data in OD matrix estimation is described and applied in a limited way at localised intersections (Ramirez, Kovacic, & Ivanjko, 2013). Number plate images were observed at the entry and exit lanes, and when matched provided the OD movements through the intersection.

Based on the potential applications for new data, service providers offer traffic surveys using the different technologies and provide “OD” matrices (BlueScan, 2016). These matrices are however measurements of trips between the locations of the recording devices. These locations are neither trip origins nor destinations when considering the full trip. Therefore, without the link between the data and the traffic model, this data may only be useful for the analysis of traffic operations in closed systems such as freeways.

Carpenter et al. (2012) used Bluetooth devices (having a similar application to ANPR) along a 15-mile section of the SR-23 in Jacksonville, Florida and the main cross roads over seven days, collecting 253 367 records from 33 789 devices. Data cleaning included ignoring records if the device was only picked up once and excluding devices that were recorded more than once at the same location due to slow moving traffic. The resultant trips were processed to record approximately 46 216 trips along the road network surveyed. A 30-minute time period was used as the maximum time to travel between sensors. The resultant trip OD matrices were developed in the same way as the expansion of traditional number plate surveys. The report concludes that the resultant OD matrix “can be used as a model validation tool when compared to a network trip table by reviewing the output from a select link analysis”. The report is silent on this analysis, but states that further research is required into the expansion factors and the optimal time for trips between sensors.

Asakura, Hato, & Kashiwadani (2000) used AVI data from the 221km tolled Han-shin Expressway covering the Osaka and Kobe areas of Japan and the system included 59 cameras in 1997. However, the study used data from a 30km section with detectors spaced between 6 and 15km creating four blocks with multiple entry and exit points related to each. Data from one five-day week (30 000 to 35 000 vehicles in 12 hours) represented 70% of the daily traffic. A linear transformation was used to expand the data using the traffic counts associated with each AVI detector. The “ODs” related to the four blocks in one direction and measured the variability

of the traffic stream on different freeway section. Their model application was limited to the corridor part of the expressway network.

Dixon & Rilett (2005) used AVI (ANPR) data sampled from the Houston Expressway and propose a methodology to expand the sample of vehicles collected from the ITS using traffic counts. They estimated the sampling rate from the traffic counts to determine the population of trips in the system and apply these to each of the AVI data at each collection site. The result was an OD matrix where the traffic zone is firstly the AVI measuring point and then extended to the freeway ramps. Using Monte Carlo simulation, it was determined that by using the OD information from AVI, potential errors for each run were reduced significantly, from 400% without the data to 36% with it. Averaging all runs reduced the error to approximately 7%.

Van Vuren & Carey (2011) evaluated the use of ANPR data for developing OD matrices and compared it to using GPS data and travel diaries. The main concern with the ANPR data is that it can be costly to achieve a high coverage and that the ANPR data is “incomplete”. On the other hand, the GPS data has a potential bias targeting only those with GPS devices. This would be a significant factor in developing countries where GPS devices are not evenly distributed throughout the entire population.

Sun, Zhu, Zhou, & Sun (2014) developed numerous quantitative metrics for tracing vehicles passing ANPR cameras deployed throughout a city road network. Their work identified different driving (route) patterns in the data for use in traffic planning and intimate that the regional distribution of the traffic patterns could be used for transport planning. The analysis of convoys (Himayounfar, Ho, Zhu, Head, & Palmer, 2011) processed 200 000 daily ANPR images. Their objective was to monitor convoys in real time to identify “unusual travel patterns”. The data was organised into clusters or convoys, which reduced the quantum of data by 98% and significantly improves processing times. It also forms the benchmarks of “normal behaviour” against which irregular or suspicious driving patterns are highlighted. Their focus was on crime prevention.

Research into the use of passive electronic survey technology has considered the use of probe data and roadside equipment (either Bluetooth or ANPR data) to assess traffic patterns and journey times and speeds. Some work has shown how ANPR data can be used to benchmark “normal traffic patterns”. Most of the work aims at improving traffic operations by feeding back into driver information to improve route choice. Limited research was found regarding the expansion of the use of ANPR (or Bluetooth) data beyond closed areas and corridors and relating the data to the wider OD trip matrices to improve traffic models to improve their base year robustness and therefore their forecasting confidence.

3.4 SUMMARY

Based on the review of the requirements of data necessary to provide optimal solutions in the estimation of OD matrices, the data needs to be current, accurate and independent while being as disaggregated as possible, i.e. related to as few OD pairs as possible. These qualities are sought in the assessment of the ANPR data. Methods to estimate OD matrices have become more sophisticated over the years, ranging from the simple bi-proportional balancing to the fusion of matrices derived from various data sources and each providing a different result. This raises the question if it is better to use the new data to find more sophisticated matrix estimation methods or use it to build on and improve current methods to provide more confidence in the results.

The work that Carpenter Fowler & Adler (2012) undertook using Bluetooth as the survey medium provided some concepts that presented a basis for the analysis of the ANPR data obtained from the GFIP freeway network.

Research into the use of passive electronic survey technology has considered the use of probe data and roadside equipment (either Bluetooth or ANPR data) to assess traffic patterns and journey times and speeds. Some work has shown how ANPR data can be used to benchmark “normal traffic patterns”. Most of the work aims at improving traffic operations by feeding back into driver information to improve route choice.

Limited research was found regarding the expansion of the use of ANPR (or Bluetooth) data beyond closed areas and corridors where the origin and destination were the measuring locations. This is limiting when relating the data to the estimation of wider OD trip matrices, which would improve traffic models in terms of their base year robustness and therefore their forecasting confidence.

4 METHODOLOGY FOR USING ANPR DATA FROM ELECTRONIC TOLL COLLECTION

This chapter describes the extent and characteristics of the ANPR data obtained from the ORT equipment deployed on the GFIP toll project. It further explains the methodology adopted to process the data into information usable in transport model development, calibration and validation.

4.1 SOURCE, EXTENT AND AVAILABILITY OF DATA

The Gauteng Freeway Improvement Project (GFIP) covers approximately 201km of the Gauteng Provinces urban freeway with gantries located at approximately 10km spacing in each direction at 42 gantries. Figure 1 shows the extent of the freeways and the location of the gantries.

The open road tolling (ORT) equipment comprises traffic surveillance and toll transaction equipment, including:

- Dedicated Short Range Communication (DSRC) equipment that communicates with the e-tags for the e-toll transactions
- Vehicle profiling equipment for the classification of each vehicle and the validation of the e-tags
- Cameras that photograph the front, rear and position of each vehicle, with Automatic Number Plate Recognition (ANPR) software to validate the e-tags and/or identify non-tagged vehicles.
- The date/time stamp of individual vehicles passing under each gantry and toll transaction.

As the gantry information is required for financial transactions and for prosecution in the event of non-payment, the project specifications called for very high levels of accuracy. Austria and Chile are countries that have successfully deployed similar ANPR systems into their ORT systems.

The gantries send the ANPR data to the Central Operations Centre (COC) in Midrand, Johannesburg, where the toll transactions are processed and the ANPR stored. SANRAL provided access to this database for the extraction of ANPR data for further analysis.



Figure 1: GFIP Freeways and Gantry Location (Source SANRAL)

The e-toll transaction information includes the vehicle classifications, which are:

- A1 – Motorcycles
- A2 – Light motor cars (<6m in length and <2.5m in height)
- B – Small heavy vehicles (>6m and <12.5m in length and/or >2.5m in height)
- C – large heavy vehicles (>12.5m in length and >2.5m in height)

4.2 ANPR DATA EXTRACTION

Large data sources used in transport planning is invariably a by-product of a system that has a non-transport function. For example, a system designed for electronic toll transactions produces ANPR data and way-finding or security systems in vehicles are trackable probes.

The e-toll team prepared a program script to extract the ANPR data in monthly text files from the main Transaction Clearing House (TCH) database. The data records contain the vehicle licence number (VLN). Replacing the VLN with a random number ID anonymises the data in terms of maintaining the privacy of private information. Each vehicle's VLN ID remains constant within each data set to ensure the potential to track vehicles through the network. The same vehicle's VLN ID may be different for different data set extractions.

Table 1 contains an example of the extent of the extracted data, indicating the total number of gantry entries per month between February 2014 and July 2015 as provided by SANRAL. In this study, the demonstration of the concepts of the use of ANPR data uses information from July 2015. Information contained in the extracted ANPR data records includes the following elements:

- The VLN ID
- The date and time the vehicle passed under the gantry
- The gantry number
- The vehicle classification

Table 1: ANPR data entries per month

Year	Month	Gantry Passes
2014	Feb	63 000 000
2014	Mar	65 766 226
2014	Apr	65 108 453
2014	May	68 607 255
2014	Jun	67 162 243
2014	Jul	72 086 282
2014	Aug	72 767 529
2014	Sep	71 694 535
2014	Oct	77 112 714
2014	Nov	73 317 825
2014	Dec	66 238 609
2015	Jan	68 163 784
2015	Feb	70 466 183
2015	Mar	78 672 333
2015	Apr	71 644 727
2015	May	75 716 814
2015	Jun	73 878 018
2015	Jul	79 407 436
Total		1 280 810 966

Appendix A provides a small sample of this data as it was obtained from the ORT system.

4.3 GANTRY SPACING

The distances between the gantries (Figure 1) are important elements in the analysis of the ANPR data. Table 2 contains the distances that vehicles would travel along logical routes along the freeway network between gantry pairs. The distances used by SANRAL to calculate the gantry tariffs were the basis for the determination of these distances. The vehicle's average speed between gantries is calculated using these distances and the times recorded as each vehicle passes under gantries. The trip length frequency distribution (TLFD) of the trips passing under two or more gantries also derived using these distances.

In Table 2, the cells shaded in blue contain the distances between consecutive gantries and the remainder of the distances are combinations of the consecutive distances. Appendix B contains the full matrix.

The gantries are located approximately 10km apart and there are generally more than one entry and exit point between them. It is therefore possible to use the freeways without passing under a gantry. There are also records of vehicles that only pass under one gantry and these vehicles will appear on the diagonals of a gantry-to-gantry matrix. For the calculation of trip length frequency distributions (Section 6.4.4), this study assumes that vehicles that only pass under one gantry travels on average 5km along the freeway. Whilst this is an estimation, it does provide a basis to measure the number of short distance trip and the same distance is applied to the ANPR data and the equivalent output from the traffic models.

Table 2: Gantry to gantry distances

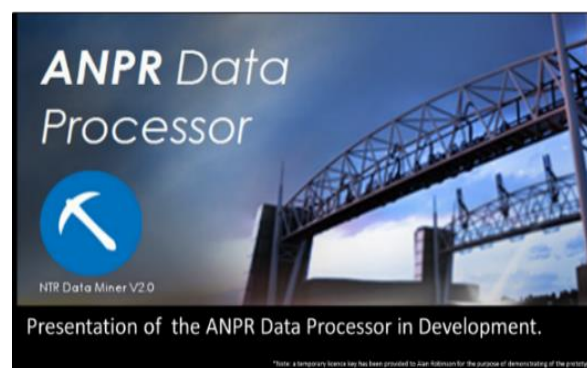
To From	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015
1001															
1002				10.0		18.6		29.9		37.2		48.2		59.5	
1003	11.5														
1004						8.6		19.9		27.2		38.2		49.5	
1005	21.1		9.6												
1006								11.3		18.6		29.6		40.9	
1007	29.9		18.4		8.8										
1008										7.3		18.3		29.6	
1009	41.1		29.6		20.0		11.3								
1010												11.0		22.3	
1011	51.3		39.8		30.2		21.4		10.2						
1012														11.3	
1013	60.8		49.3		39.6		30.9		19.6		9.5				
1014															
1015	69.3		57.8		48.2		39.4		28.2		18.0		8.6		

Notes:

1. 1001 corresponds to Gantry 1 as displayed in Figure 1.
2. This table only displays a portion of the overall matrix for clarity. Appendix A contains the full matrix.

4.4 PROCESSING THE ANPR DATA

Standard database programs cannot be used due to the size of the ANPR data files, therefore bespoke software was developed to process the ANPR data¹. The ANPR data was provided in monthly files each in the order of 2.4Gb. The monthly data was processed for two purposes:



- Deriving traffic counts from the gantry data, and
- Preparing and extracting Gantry to Gantry trip information

An overview of the above process is provided below and more detailed description of the Data Miner software and the processing of the ANPR data is provided in Appendix C.

4.4.1 Extracting Traffic Counts

The ANPR data was sorted by gantry, vehicle class and time. The numbers of vehicles in each gantry and vehicle class set were counted in 15 minute or hourly bins. Table 3 provides an example of the average weekday count at gantry 8, southbound on the N1 between the Allandale and Buccleuch Interchanges.

The ANPR Data Miner software permits the selection of time, day or days and vehicle type. Appendix D contains gantry counts for all other gantries.

Table 3: Example of Counts from Gantries

Gantry	8			
	Light	Sm Hgv	Lg Hgv	Total
0	319	28	48	394
1	178	36	54	268
2	133	32	46	211
3	179	32	40	251
4	624	51	56	731
5	4515	141	94	4749
6	11448	221	89	11758
7	10923	257	95	11275
8	8726	373	117	9216
9	7635	453	148	8236
10	6702	428	158	7289
11	6562	434	157	7154
12	6691	462	160	7313
13	6700	469	171	7340
14	6548	480	162	7190
15	7709	462	155	8327
16	8494	386	143	9023
17	7378	307	137	7822
18	5100	186	120	5406
19	3309	120	94	3522
20	2256	86	84	2426
21	1662	69	79	1810
22	1168	53	61	1282
23	673	39	47	758
Total	115631	5605	2513	123750

¹ Data Miner was developed specifically for the processing of the ANPR data and used with permission of the program developer

4.4.2 ANPR Data accuracy

In this study, the hypothesis is that ANPR data is more accurate and ultimately more comprehensive and reliable than traditional data sources and comparable to, if not better than, other new data sources that fall under the “big data” banner. In this section, comparative data collection methods are reviewed and compared to the ANPR data obtained for this study. The evaluation of the ANPR data accuracy comprises:

- Comparing ANPR data to traditional traffic survey methods
- Assessing error sources in the ANPR
- Comparing the ANPR data to other “big data”.

4.4.2.1 Comparison of ANPR data to continuous electronic count data

Throughout the GFIP freeway network, permanent electronic traffic counters are installed on interchange approaches as part of the project. The information from these counters are included in the SANRAL’s Comprehensive Traffic Observation (CTO) database. At the time of preparing this study, the 2015 CTO data had not been released; however, the 2014 CTO data and the 2014 ANPR data had been obtained. The comparison of the CTO and ANPR counts was undertaken for February and October 2014. Both sets of counts were binned into 15-minute time intervals and the average flow profiles were prepared for the average weekday, the average weekend and the average for the 7-day week. If there were time periods where there were missing counts, a zero was entered for the count and having the effect of reducing the 15-minute average. The list of gantry and associated CTO station numbers are provided in Appendix E.

The CTO counters are electronic loop counters located in each lane of the on-ramps and main through lanes of interchanges. The sum of the lanes provides the total traffic volumes that would pass under a downstream toll gantry. Any faults with the loops in any one lane or with the loggers will impact on the counts. These counts are undertaken in terms of COTO TMH 14. (COTO, 2013).

The graphs in **Error! Reference source not found.** below show the comparison of the two data sets. The CTO counts are often slightly below the ANPR data.

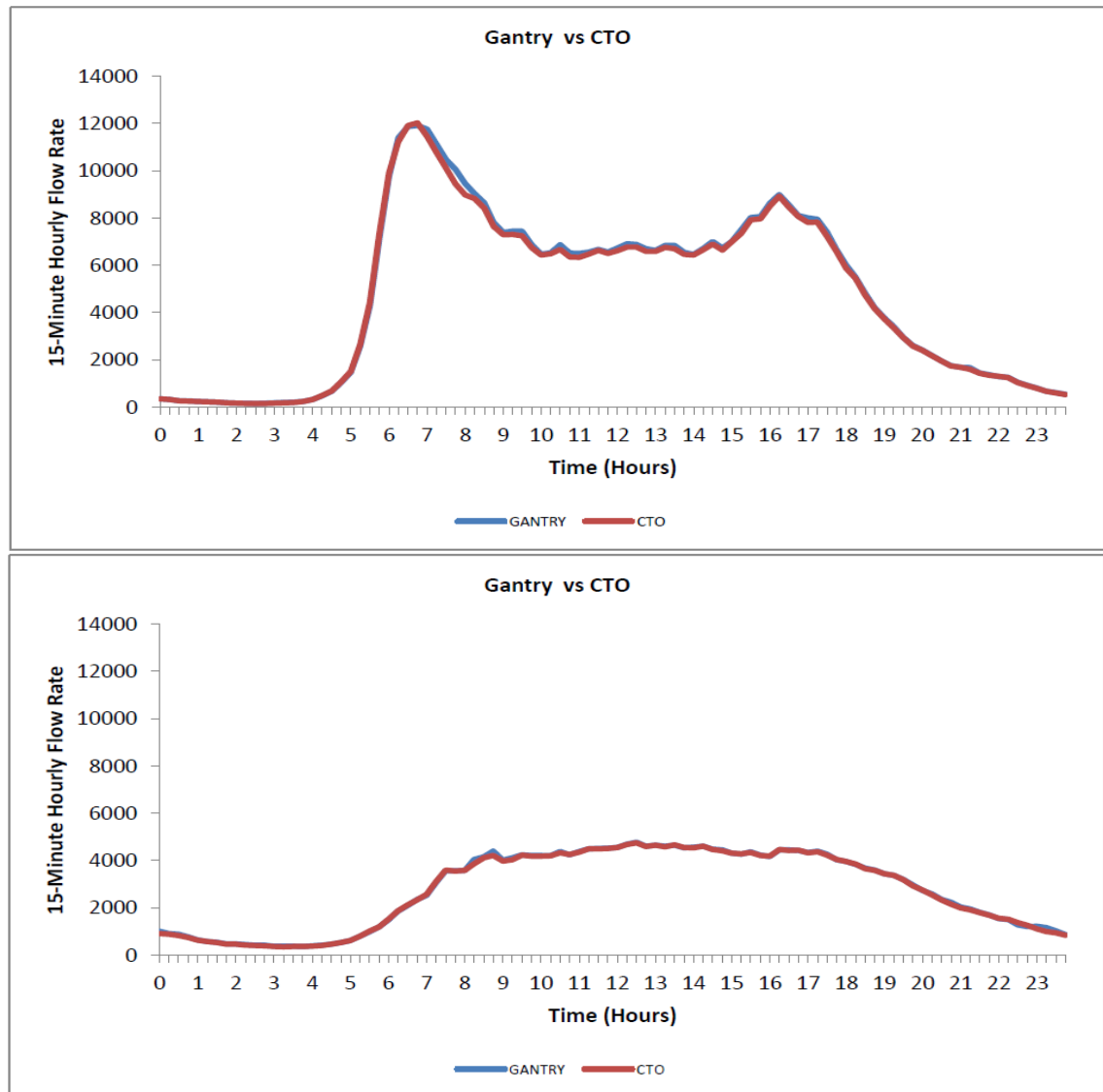


Figure 2: Comparison of ANRP (Gantry 8) and CTO Data (Station 1863)

Additional comparisons are provided in Appendix F.

A detailed comparison of the two data sets highlighted that discrepancies were commonly a result of missing entries in the CTO data. Most of the ANPR/CTO comparisons showed minor discrepancies with missing or slightly lower counts during peak often occurring due to stationary vehicles on the induction loops. Larger discrepancies occur when there are equipment failures in the CTO recording equipment resulting in extended periods where there is no CTO data. Where there is no data, zeros are entered in raw data provided and the calculated averages included these zeros which reduced the average CTO counts.

These comparisons indicate that the ANPR count data is more complete and can therefore be considered more comprehensive and more accurate than the CTO data. The CTO data remains a trusted and comprehensive traffic count database and currently used as the benchmark for traffic modelling studies in South Africa. It must be noted that the calculation of the published CTO data makes informed adjustments when dealing with the missing data, a process not required in the ANPR data.

ANPR data is a combination of traffic count data at each gantry location and has the option to track vehicles between gantries, thus providing continuous information on the distribution of traffic through the freeway network.

Adding the ORT vehicle classification functionality, the continuous data from the e-toll system provides one of the most comprehensive traffic data sets available.

4.4.2.2 ANPR data sources of error

The e-toll equipment is required to provide comprehensive and accurate data as the information from each transaction is used for financial transactions and needs to be accurate enough to comply with the South African Legal Metrology Act (Act 9 of 2014)². However, some issues that can detract from the accuracy of the data include:

- In South Africa, amongst other issues, motorists are permitted to use temporary licence plates under certain circumstances, for example when buying a new car or re-registering one from another province, one is permitted to use a paper registration placed anywhere in the rear window for a period of 21 days. As a result, these vehicles do not display number plates and cannot be recognised by the ANPR system. All vehicles without number plates are given the same VLN ID during the anonymising of the data.
- There are also instances of number plate duplication, where more than one vehicle has the same number plate. These are picked up during a checking algorithm to identify illogical matching, i.e. where a number plate is recorded at very different places within seconds of each other and then again minutes later indicating more than one vehicle. Identified inconsistent or cloned number plates are all given a second VLN ID.
- A third instance where a repeated VLN ID is used is in unreadable number plates which may be damaged, obscured or dirty. In the data set, these three VLN IDs were recorded and included in the derivation of the gantry counts but excluded from the gantry to gantry count matrices.

² Government Gazette No. 37661. 19 May 2014 – Legal Metrology Act 2014, Republic of South Africa

The number of occurrences of the above in the sample data are provided in the **Error! Reference source not found.** It should be noted that the above would only affect the distribution data and not the count data, which is not dependent on “number plate” matching.

Table 4: ANPR Data Errors

Record Description	Number of Entries	Percentage of Sample
No Number-plate	1,472,409	1.9%
Unreadable/damaged number	1,108,400	1.4%
Illogical movements	1,061,170	1.3%
Total records not used for trips	3,641,979	4.6%
Total Gantry passes	79,407,436	100.0%

4.4.3 Comparison to other big data sources

The collection of transport related “big data” has focused on probe data methods, i.e. GSM (cell phone) tracking and blue tooth tracking. In addition to this, there is other GPS probe data coming from satellite tracking and in-vehicle security devices. Such data sources include Tomtom™ and Tracker™ data.

This section briefly compares these alternative data sources with the ANPR data in terms of their potential accuracy, coverage and use in the development, calibration and validation of strategic traffic models.

The cellular service providers can provide GSM (cell phone tracking) data. Unfortunately, the ability to access this data is somewhat difficult, predominantly due to stated privacy issues.

The advantages of probe data include:

- A very wide area of coverage
- The ability to track individual trips from one geographical location to another.
- An immense database of trip data on a 24-hour basis.

The negatives associated with this data includes the following:

- The ability to distinguish the relationship between devices and travelling persons or vehicles
- The accuracy of the origin-destination and route information when the GSM information is obtained from the communication towers

- The ability to determine the expansion factors to calculate the sample size and the predicted travel demand
- The potential bias in the data since these devices are not evenly spread through all socio-economic groups

Bluetooth tracking is a similar concept to ANPR data as the Bluetooth devices are recorded passing a specific reader. This provides a similar type of database as the ANPR data except that the device is not necessarily linked to a vehicle. The actual sample size is unknown making the expansion of the data to traffic volumes and, like the probe data, the potential for bias in the data towards specific population groups will be difficult to determine.

4.5 GANTRY TO GANTRY DATA

Each gantry in the network has a unique number and location for identification purposes. Gantry to gantry (G2G) data are displayed as matrices where each cell represents the occurrence or attribute of traffic movements between one gantry and another. These attributes can be the number of trips, journey times or speeds, or distances between the gantries.

4.5.1 Extraction of trip data

In traffic modelling terms, a trip is defined as the movement of a vehicle through the road network from its origin to its destination. The G2G portion of a trip is the portion of the whole trip where a vehicle passes under one or more gantries by travelling along the freeway.

The definition of a G2G trip is therefore when a vehicle is recorded at one gantry and then either:

- Does not pass another gantry,
- Does not pass another gantry in the specified time interval, or
- Passes successive gantries with each pass taking place within the specified time interval

The time interval is assumed to be a reasonable time period for a vehicle to travel between successive gantries. If a vehicle exceeds this time interval, the assumption is that the vehicle would have left the freeway to fulfil a specific purpose of this part of the trip, e.g. drop off children at school. This is assumed to represent the completion of that trip. If the vehicle is recorded at the next logical gantry after the last record but after the allowable time, this is considered the first entry of a new trip e.g. from the school to the next destination.

This time interval would change by time of day. During the peak hours, there are sections of the freeway that travel at speeds below 30kph and a vehicle may take over 20 minutes to travel between gantries. During off-peak periods, this may be 10 minutes travelling at 60kph. This

variability was built into the software functionality when the raw gantry entries are processed into individual trips. (refer to Section 4.5.1)

With the ANPR data, only tracking vehicles on the freeways, the actual trip origins and destination are unknown. However, the proximity of the entry and exit points can be identified and there is a limited number of origin-destination pairs that would use this combination of entry and exit points. This group of origin-destination pairs is significantly less than those that would pass through a single point traffic count.

The G2G trips are presented in a matrix format with each cell of the matrix providing the number of vehicles that pass between specific combinations of gantries and the matrix cell indicating the first and last gantry passed. This process and the outcomes are described in Section 4.5.2 below.

Extracting the G2G trip information incorporated the following processes:

1. Sort the data by VLN ID (vehicle) and by time. This produces a list of gantries that each vehicle passes in chronological order.
2. Record the first gantry number that initially identified each vehicle.
3. Read the next entry and check that the gantry is a logical successor to the previous gantry using the G2G distance matrix and if so calculate the time interval between the previous entry and this entry.
4. Read the next entry for the same vehicle and if the time lapse is within the specified maximum time and the gantry is a logical successor to the previous entry based on the G2G distance matrix.
5. Repeat step 4 until the next entry has a gantry that is not a logical successor to the previous record or the time interval is exceeded. The previous entry then contains the last gantry number of the trip and the time the vehicle passed this gantry. The current entry is recorded as the start of a new trip.
6. If the next entry relates to a different vehicle, the previous entry is recorded as the last gantry passed in the previous vehicle's trip. The current entry is recorded as the start of a new trip.

A set of G2G trip records is created with each containing the following information:

- The vehicle ID
- First gantry number of the trip
- The date and time the first gantry was passed

- Last gantry number of the trip
- The date and time the last gantry was passed
- Vehicle classification

4.5.2 Gantry to gantry count matrices

G2G trip records contain the first and last gantry numbers and the date and time associated with the first and last gantry passed. A single gantry pass would be on the diagonal of the G2G count matrix³.

TO FROM	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020
1001	1629																			
1002		3867		1296		710		640		95		53		13		20				36
1003	883		2769																	
1004				530		219		191		31		17		4		4				17
1005	237		702		1573															
1006						2360		1790		248		134		32		36				198
1007	236		578		1321		939													
1008								3244		775		334		71		51				524
1009	52		148		258		188		1601											625
1010										727		391		66		61				0
1011	22		44		82		118		744		1138									271
1012												757		386		378				
1013	11		17		42		52		378		867		854							95
1014														913		966				
1015	6		11		26		21		92		342		714		1113					2
1016																560				
1017	10		13		29		19		82		244		588		1076		2169	2	5	13
1018	19		36		80		89			317		102		11		1		461		
1019																				488
1020	18		31		118		162			577		196		9		1		1179	34	1059

Figure 3 provides a sample G2G trip matrix for all vehicles travelling between gantries 1 to 20 (refer to Figure 1) between 07:00 and 08:00 on weekday mornings.

³ For presentation purposes, only the first 20 gantries are show in the text. The Appendix D contains the full matrices of the counts, times and speeds.

TO FROM	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020
1001	1629																			
1002		3867		1296		710		640		95		53		13		20			36	
1003	883		2769																	
1004				530		219		191		31		17		4		4			17	
1005	237		702		1573															
1006						2360		1790		248		134		32		36				198
1007	236		578		1321		939													
1008								3244		775		334		71		51				524
1009	52		148		258		188		1601											625
1010										727		391		66		61				0
1011	22		44		82		118		744		1138									271
1012												757		386		378				
1013	11		17		42		52		378		867		854							95
1014														913		966				
1015	6		11		26		21		92		342		714		1113					2
1016																560				
1017	10		13		29		19		82		244		588		1076		2169	2	5	13
1018	19		36		80		89			317		102		11		1			461	
1019																				488
1020	18		31		118		162			577		196		9		1		1179	34	1059

Figure 3: Example of the average G2G counts for all vehicles on weekdays between 07:00 and 08:00

Electronic traffic counts provide comprehensive classified traffic information at one location in the road network. A single ANPR camera would provide similar information as shown above. However, two electronic counts provide two counts while two ANPR cameras provide three counts; two for vehicles that only pass under each of the cameras and one that passes under both cameras. The potential number of G2G counts from a series of N gantries is a triangular series where:

$$\text{G2G counts} = N(N+1)/2$$

Theoretically, if all 42 gantries on the freeway network in a straight line, they would produce 903 G2G counts. The fact that the ANPR gantries are partially on a circular route (N1, N3 and N12) and on parallel routes (N1 and R21), the number of logical combinations are reduced to 340 counts. With more cameras, the amount of data increases significantly, which would then make this means of data collection more economical.

4.5.3 Gantry to gantry journey times

The record for each trip contains the first and last gantry passed and the date/time stamps. The time taken to travel from one gantry to another is therefore calculated and presented in a matrix format. Figure 4 provides a sample of the average journey times, in minutes, to travel between gantries between 07:00 and 08:00 on weekday mornings.

Times (min) from 7.00 to 8.00

To Gantry From Gantry	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	
1001																					
1002				7.92		21		32.1		38.4		49.1		57.1		61.2				37.7	
1003	7.55																				
1004						14.7		26.1		32.2		45.2		50.6		55.6				30.9	
1005	13.6		5.767																		
1006								12		18.8		29.8		37.4		42				17.4	
1007	18.9		10.88		5.2																
1008										6.13		16.4		23.8		27.8				5.08	
1009	26.5		18.03		12.4		7.72													4.53	
1010												9.78		16.8		20.6					
1011	45.2		35.47		30.1		24.9		17.2											21.7	
1012														6.68		10.8					
1013	59.6		50.4		43.8		38.9		32.1		15.3									37.4	
1014																4.35				40.4	
1015	65.2		57.82		52.1		46.7		39.7		23.2		8.77		0.08					46.3	
1016																				40.2	
1017	76.9		68.98		63		57.9		49		33.4		17.2		9.63				47	55	42.4
1018	29.7		20.42		14.3		8.93			5.35		15.8		23.4		26.4					
1019																					
1020	34.4		25.18		19.3		14.2			11.1		21.5		28.6		40			5.63		

Figure 4: Example of an average G2G travel time matrix on weekdays between 07:00 and 08:00

The G2G journey time for the trips recorded at only one gantry is zero as there is no second gantry record and time stamp.

4.5.4 Gantry to gantry speeds

The G2G average speeds are calculated from the distances between gantries and the average journey times of vehicle passing under the corresponding gantries. As there is zero distance associated with a single gantry pass, the average speed for these trips not estimated. The data is used for the validation of the network where vehicles pass under numerous gantries as they travel the lengths of freeway.

4.6 SUMMARY

The requirements for the development of OD matrices, as described in Section 3, is to have accurate, independent and disaggregated traffic data. The assessment of the ANPR data has shown that it goes a long way towards meeting these needs. In summary, the ANPR data that was provided for this study offers the following:

- Long term independent traffic information
- The data is accurate when compared to other data sources
- For the use of this data in matrix estimation, it can be disaggregated into individual and independent counts that are related to fewer OD pairs.

It is also reported that ANPR installations are relatively expensive (van Vuren & Carey, 2011), however this work has shown that the number of counts over distance increases according to a triangular series, which makes ANPR more economical with the addition of cameras deployed.

5 DEVELOPMENT OF A METHODOLOGY FOR MODEL VALIDATION USING ANPR DATA

The gantry to gantry (G2G) count matrices derived from the ANPR data represent the average number of trips between ANPR sites for the days and hours selected to represent the modelled time period, for example, the average trips made on weekdays during the morning peak period. Currently, because the G2G counts data is derived from locations that do not define a closed cordon, equivalent matrices are not readily derived from the traffic model. This section describes the methodology to derive a formula to calculate the equivalent G2G count matrices from a traffic model. The “counts” in the G2G matrices from the traffic model will be from the assigned trips on the links with the gantry pairs.

The formula development utilises select link (SL) trip matrices, which are derived from a standard modelling process. The select link matrix contains the proportion of the modelled trip matrix that use the selected link, i.e. the $T_{ij} \cdot P_{ija}$ in equation 2 above and described in Section 3.1. A stepwise process is described, whereby the trips in select link trip matrices are assessed and combined to isolate the trips in cells that result in the combination of the select link matrices associated with any two gantry locations. Combining all possible gantry location pairs results in the equivalent modelled G2G matrix. Therefore, the comparison of the ANPR and modelled G2G matrices is another means to validate the traffic model.

5.1 G2G FORMULA DEVELOPMENT

The G2G count matrices are based on the trips that pass under each gantry and series of gantries. As with the calibration of a trip matrix from counts, the proportion of the trips in the OD matrix that pass through this location must be identified. Extracting a SL trip matrix from the link with the ANPR reader (gantry) achieves this and combining these SL matrices in a logical way will provide an equivalent SL to SL matrix for validation against the G2G matrix.

A SL matrix contains the ODs of the trips that pass along a selected section of road or model link. However, each select link matrix contains trips that pass through one or more upstream gantries and/or one or more downstream gantries. Figure 5 demonstrates the possible trips recorded through a freeway section with a gantry “B” and numbered entry/exit points or zones numbered 1 to 14.

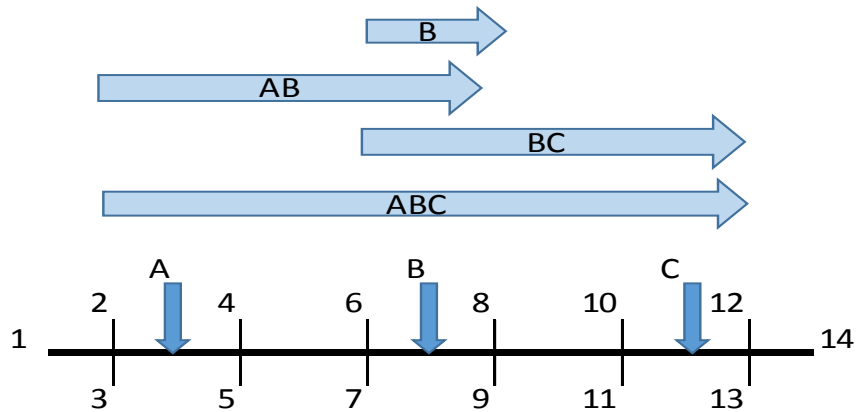


Figure 5: Select Link Trips Passing Gantry B

The SL matrix would be the full 14 by 14 trip matrix but the cells would only contain trips with origins and destinations as indicated by the arrows. For example:

- Only B would include trips from origin zones 4 to 7 to destination zones 8 to 11
- AB would include trips from origin zones 1 to 3 to destination zones 8 to 11
- BC would include trips from origin zones 4 to 7 to destination zones 12 to 14
- ABC would include trips from origin zones 1 to 3 to destination zones 12 to 14

The SL matrix for gantry B would contain all the above trips but the G2G matrix for trips that only pass through B only include the first element. The formula uses SL matrices from all gantries to isolate only those trips that pass gantry B

In matrix format, the cells that could contain trips through gantry A are marked in the matrix with “A”, i.e trips entering the network from zones 1, 2 and 3 and exit through zones 4 to 14. This includes trips that pass gantry A only as well as through gantry A and B, and gantries A, B and C. This would be Select Link A (SL_A) as shown in Matrix 1.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1				A	A	A	A	A	A	A	A	A	A	A
2				A	A	A	A	A	A	A	A	A	A	A
3				A	A	A	A	A	A	A	A	A	A	A
4														
5														
6														
7														
8														
9														
10														
11														
12														
13														
14														

Matrix 1: Select Link from location A

Similarly, the cells that would contain trips that are included in a select link matrix through gantry B are marked with “B”. This includes trips entering at zones 1 to 7 and exiting through zones 8 to 14. These trips include trips that pass gantry A and B, only gantry B, gantries B and C and through gantries A, B and C. This would produce the Select Link B (SL_B), as in Matrix 2.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1								B	B	B	B	B	B	B
2								B	B	B	B	B	B	B
3								B	B	B	B	B	B	B
4								B	B	B	B	B	B	B
5								B	B	B	B	B	B	B
6								B	B	B	B	B	B	B
7								B	B	B	B	B	B	B
8														
9														
10														
11														
12														
13														
14														

Matrix 2: Select Link from location B

Likewise, cells that contain trips that are included in a select link matrix through gantry C are marked with “C”. This included trips entering through zones 1 to 11 and exiting through zones 12 to 14. These trips include trips that pass gantry A, B and C, B and C and C only. This would be Select Link C (SL_C), as in Matrix 3

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1												C	C	C
2												C	C	C
3												C	C	C
4												C	C	C
5												C	C	C
6												C	C	C
7												C	C	C
8												C	C	C
9												C	C	C
10												C	C	C
11												C	C	C
12														
13														
14														

Matrix 3: Select Link from location C

Combining the three Select Link matrices to include all the trips that would go through the three gantry locations will populate the cells as shown with A, B, C, AB, BC and ABC. (SL_{all}) as shown in Matrix 4

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1				A	A	A	A	AB	AB	AB	AB	ABC	ABC	ABC
2				A	A	A	A	AB	AB	AB	AB	ABC	ABC	ABC
3				A	A	A	A	AB	AB	AB	AB	ABC	ABC	ABC
4								B	B	B	B	BC	BC	BC
5								B	B	B	B	BC	BC	BC
6								B	B	B	B	BC	BC	BC
7								B	B	B	B	BC	BC	BC
8												C	C	C
9												C	C	C
10												C	C	C
11												C	C	C
12														
13														
14														

Matrix 4: Select Link from location A, B and C

Therefore, any combination of gantries is isolated by the combination of letters in the cells. Trips that only pass through gantry A and not the others only contain the letter A. Trips that pass gantry A and B and not C, are contained in the cells marked with AB and so on.

Whilst it would appear a relatively simple process to use conditional statements to isolate the relevant cells of any combination of gantry pairs, when there are over 40 gantries with >350 logical G2G pairing combinations and a matrix with 1000 zones, the task becomes a little more challenging. Furthermore, extending this process where any number of ANPR camera can be set up throughout a city can be used for data collection and model validation.

Examining the matrices for SL_A , SL_B , SL_C and SL_{all} , for a single cell, a trip that passes through gantry A, B and C in the three individual matrices and the combined matrix are the same. If one is to assume that in this example, exactly one trip travels from one zone to each of the others, meaning that 13 trips enter and exit at each zone, each cell of the matrix (with the exception of the diagonal cells would contain a “1”).

Therefore:

$$\begin{aligned}
 \text{Where cells that contain:} \quad & AB: & A = B = AB \\
 \text{Where cells that contain:} \quad & BC: & B = C = BC \\
 \text{Where cells that contain:} \quad & ABC: & A = B = C = ABC
 \end{aligned}$$

Equation 4

With this insight into the contents of the cells of the select link matrices, it is possible to derive a methodology and a formula to isolate the trips that only pass between one or more locations applicable to any trip matrix from any sized model.

5.2 MODELLED PARTIAL OD TO G2G FORMULA

The first step in developing a formula that will isolate the cells of an OD matrix whose trips pass through two locations A and B and no other up- or downstream locations is to consider two of the select link matrices. SL_A and SL_B .

Combining these two SL matrices provides the trips in the cells in Matrix 5. However, the cells of interest are only those that contain the letters AB.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1				A	A	A	A	AB	AB	AB	AB	AB	AB	AB
2				A	A	A	A	AB	AB	AB	AB	AB	AB	AB
3				A	A	A	A	AB	AB	AB	AB	AB	AB	AB
4								B	B	B	B	B	B	B
5								B	B	B	B	B	B	B
6								B	B	B	B	B	B	B
7								B	B	B	B	B	B	B
8														
9														
10														
11														
12														
13														
14														

Matrix 5: Select Link from location A and B

Using the Hadamard product (Horn & Johnson, 2012) on SL_A and SL_B will produce a zero where there is only an A or a B and an $AB^2 = A^2 = B^2$ (Equation 4) in the cells containing the AB. The square root of the result will produce the matrix with all the cells that contain AB.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1								AB	AB	AB	AB	ABC	ABC	ABC
2								AB	AB	AB	AB	ABC	ABC	ABC
3								AB	AB	AB	AB	ABC	ABC	ABC
4														
5														
6														
7														
8														
9														
10														
11														
12														
13														
14														

Matrix 6: Step 1 result with cells AB and ABC

This will also contain the cells containing ABC with the result provided in Matrix 6.

Note that the Hadamard product is used throughout this process, which is the multiplication of corresponding cells as opposed to the standard matrix multiplication.

The sum of the trips that pass gantry location A and B is therefore:

$$T_{G_{AB}} = \sum_{ij} \{ \sqrt{SL_a \cdot SL_b} \}$$

Equation 5

Where

- $T_{G_{AB}}$ = the trips through gantry location A and B
- SL_a = Select Link Matrix through gantry location A
- SL_b = Select Link Matrix through gantry location B

Removing the trips in the cells containing ABC will leave the trips that only pass through location A and B without passing through C. Equation 4 shows that values in the cells of the select link matrices, and the combination of select link matrices, in the cells bound by rows 1⇒3 and columns 12⇒14, are equal.

Therefore, subtracting the select link from location C from the matrix in Matrix 6 will result in zero in the cells that contain ABC but -C in rows 4⇒14 and columns 12⇒14. The result is Matrix 7.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1								AB	AB	AB	AB			
2								AB	AB	AB	AB			
3								AB	AB	AB	AB			
4												-C	-C	-C
5												-C	-C	-C
6												-C	-C	-C
7												-C	-C	-C
8												-C	-C	-C
9												-C	-C	-C
10												-C	-C	-C
11												-C	-C	-C
12												-C	-C	-C
13												-C	-C	-C
14												-C	-C	-C

Matrix 7: Step 1 less upstream SL_{B+1}

Taking the positive cells of Matrix 7 provides the desired result of the cells that contain only the trips that pass through only gantry A and B as provided in Matrix 8.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1								AB	AB	AB	AB			
2								AB	AB	AB	AB			
3								AB	AB	AB	AB			
4														
5														
6														
7														
8														
9														
10														
11														
12														
13														
14														

Matrix 8: Resultant Matrix of Trips through Only Location A and B

Similarly, isolating the cells containing the trips that only pass through gantry locations B is possible. Equation 5 does not change the SL_B matrix since both input matrices are SL_B . (Matrix 2) However the result includes trips through A (AB), C (BC) and A and C (ABC) as shown in Matrix 9.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1								AB	AB	AB	AB	ABC	ABC	ABC
2								AB	AB	AB	AB	ABC	ABC	ABC
3								AB	AB	AB	AB	ABC	ABC	ABC
4								B	B	B	B	BC	BC	BC
5								B	B	B	B	BC	BC	BC
6								B	B	B	B	BC	BC	BC
7								B	B	B	B	BC	BC	BC
8														
9														
10														
11														
12														
13														
14														

Matrix 9: Trips through gantry B including those through A and C

Subtracting SL_A and SL_C will reduce the cells containing AB, BC and ABC. These are the gantry locations upstream and downstream of the (two) gantries chosen for analysis. Therefore, subtracting SL_A and SL_C will result in the cells containing AB, BC and being zero but there will be a -A, -C and -A-C in the overlapping cells as shown in Matrix 10. The positive values will result in a matrix that only contains B in rows 4⇒7 and columns 8⇒11.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1				-A	-A	-A	-A					-A-C	-A-C	-A-C
2				-A	-A	-A	-A					-A-C	-A-C	-A-C
3				-A	-A	-A	-A					-A-C	-A-C	-A-C
4								B	B	B	B			
5								B	B	B	B			
6								B	B	B	B			
7								B	B	B	B			
8												-C	-C	-C
9												-C	-C	-C
10												-C	-C	-C
11												-C	-C	-C
12												-C	-C	-C
13												-C	-C	-C
14												-C	-C	-C

Matrix 10: Removal of trips through gantry A and gantry B

Identify the cells that contain trips that pass through all upstream and downstream gantry locations is therefore achieved using:

$$T_{(a-1)+(b+1)} = \sum SL_{a-1} + \sum SL_{b+1}$$

Equation 6

Where:

$T_{(a-1)+(b+1)}$ = the trips that pass through the gantries upstream and downstream of the subject gantries

SL_{a-1} = the select link matrix through the gantry upstream of gantry A

SL_{b+1} = the select link matrix through the gantry downstream of gantry B

In each instance, the sum of the isolated positive cells provides the equivalent G2G count from the model that can be compared to the actual G2G counts from the ANPR data. Equation 7 represents this relationship.

$$T_{G_{ab}} = \sum_{T_{ij}>0} \left\{ \sqrt{SL_a \cdot SL_b} - \left(\sum SL_{a-1} + \sum SL_{b+1} \right) \right\}$$

Equation 7

Where:

$T_{G_{ab}}$ = Trips from Gantry(a) to Gantry(b) only

SL_a = Select link matrix through Gantry(a)

SL_b = Select link matrix through Gantry(b)

$\sum SL_{a-1}$ = Select link matrix(ies) of gantry(ies) upstream of Gantry(a)

$\sum SL_{b+1}$ = Select link matrix(ies) of gantry(ies) downstream of Gantry(b)

5.3 FORMULA TESTING

The South African National Roads Agency (SANRAL) granted permission to use the Gauteng Freeway Improvement Project (GFIP) traffic model to test the viability of the above formula. Further information on the use of the model to prepare the comparative model outputs is provided in Appendix G.

Extracting select link trip matrices from each link on which a gantry is located (ANPR site) provided the basis for testing the formula.

The objective of testing the formula is to ensure that the equivalent G2G matrix derived from the model will reproduce the expected trip pattern. This means, when assigning the resultant select link to select link (SL2SL) matrix to the network, it should include trips entering the network between the first gantry location and the one preceding it, and leaving the network between the second gantry and the one after it. However, for proof of concept testing, using the full model is not practical.

Deriving the SL matrices from a loaded model assignment will result in them containing trips for which the freeway is the primary route choice and trips that use the freeway only because of congested conditions on the primary route being on local roads. The matrices are checked by assigning them to an empty network. All trips would be assigned to their primary route choice, which may not necessarily be on the freeway between the selected gantries.

Therefore, for demonstration purposes a “trimmed” GFIP model was created to ensure that trips recorded on the freeways stay on the freeways. Using the “Cordon” function (SATCH) in the SATURN program produced the cordon network and the full set of six stacked user-class trip matrices contained in this model. The six user classes were added to create a single user class matrix.

Generating the equivalent G2G trip matrices required the following steps using the SATRUN software:

- Develop a batch file to generate select link matrices from all gantry locations
- Use the G2G distance matrix to determine logical G2G combinations
- Prepare an association table that identifies all the up- and downstream gantry locations
- Develop a batch file that calls the required SL matrices to calculate the SL2SL trip matrix
- Assign the SL2SL trip matrix to the freeway network to ensure that trips only travel between the designated gantry locations

5.3.1 Freeway network model

Figure 6 depicts the cordoned model network extracted from the full GFIP model and shows the freeways and a traffic assignment (blue bands).

The test used to ensure that G2G equivalent matrices derived from this model using the above formula would mean that, when assigned to this network:

- Trips would enter the network only at interchanges between the first selected gantry location and the previous one.
- Trips would leave the network at interchanges after the second selected gantry and not pass through a downstream gantry location
- No trips should enter or leave the freeway at any interchange between the two selected gantry locations, and
- The number of trips on all freeway links between the selected gantry locations should be equal.

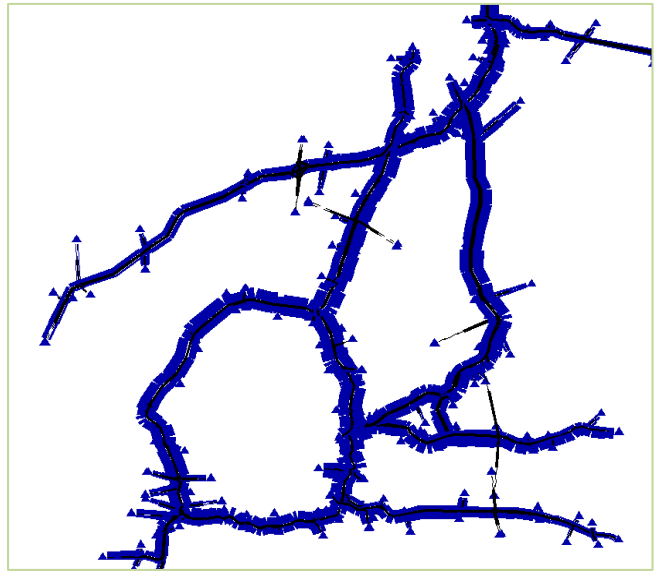


Figure 6: Modelled Freeways Assigned AM Peak Hour 2015

The following sub-sections present examples of the results obtained from these tests.

5.3.2 Preparation of Select Link (SL) matrices

A total of 42 SL matrices were extracted from the freeway model using a batch file. Each matrix contained all the trips that passed the gantry of the corresponding number.

5.3.3 Identification of gantry associations

Reference is made to upstream and downstream gantries associated with the two gantries for which the G2G count is sought in Equation 7. The SL matrices at these associated gantries are used to prevent trips going beyond these points in the network. A table of upstream and downstream gantry associations was prepared and was used in the automation of the use of the formula in a batch file. A sample from the full table is provided in Table 5. The numbers refer to the gantry numbers in Figure 1.

Table 5: Example of the Association Table of Up- and Down Stream Gantries

Gantry	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Upstream	1	3	0	5	2	7	4	9	6	11	8	13	10	15	12	17	14	0	20	8	22
	2	37	0	0	37	0	0	18	0	0	18	0	0	0	0	28	0	0	0	10	32
	3	38	0	0	38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Downstream	1	0	4	1	6	3	8	5	10	7	12	9	14	11	16	13	29	15	7	21	18
	2	0	40	40	0	0	0	0	19	19	0	0	0	0	0	0	0	29	10	0	0
	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

The contents of the table provide the upstream and downstream gantries to the gantries listed across the top of the table. For example:

- Gantry 1 has gantries 3, 37 and 38 feeding traffic towards gantry 1 while there are no downstream gantries
- Gantry 4 has gantries 2, 37 and 38 upstream and gantry 6 downstream.

5.3.4 G2G Matrices

The G2G distance matrix (Table 2 in Section 4.3) was used to identify gantry combinations that have non-zero distances. The row and column numbers represent the first and last gantries for which the G2G and equivalent SL2SL trip matrices were required. The matrix diagonals represent trips that only pass through the one gantry and although the distance associated with a single gantry pass is zero, they were included in the process.

The SL2SL matrices were prepared by calling the required select link matrices into the SATURN matrix calculation module and a “key” file containing the procedures, including Equation 7 in Section 5.2. The resultant matrices are full sized model matrices containing the trips passing through the selected gantries and therefore providing the modelled origins and destinations of these trips.

The total number of trips in these SL2SL matrices provide the volumes in the cells of the G2G matrices that are the counts over distance.

5.3.5 Assignment of SL2SL matrices

The resultant SL2SL trip matrices from the traffic model are to represent the trips that only pass through and between the selected gantries, i.e. a G2G count over distance. Assigning these matrices to the network results in an assignment of trips along the selected section of freeway. Note that there must be no trips that pass through the gantries upstream or downstream of the selected gantries and no trips can enter or leave the freeway between the selected gantries. The following example is provided to demonstrate this result.

In this example, Figure 7 shows the assignment of trips between gantry 13 (between 14th Avenue and Beyers Naude Interchanges) and gantry 7 (between New Road and Olifantsfontein Interchange). These trips enter the freeway between gantry 11 and 13, travel north on the N1 and exit the freeway between gantry 7 and 5. No trips enter or leave the N1 between gantry 13 and 7, including to the M1 or N3.

In this example there are no competing routes as the model used for this demonstration is only of the freeway network. To ensure that the calculation of the G2G model matrices is robust other examples were tested.

An example that does offer alternative routes is between gantry 8 and 23, southbound trips between Midrand (from the Olifantsfontein, New Road and Allandale Interchanges to Reading (R59) Interchange on the N12 and the Old Barn Interchange on the N3)

In this instance, the assignment shows trips from the N1 north of the Olifantsfontein Interchange as shown in Figure 8. This is contrary to the intention of the formula and the reason was investigated.

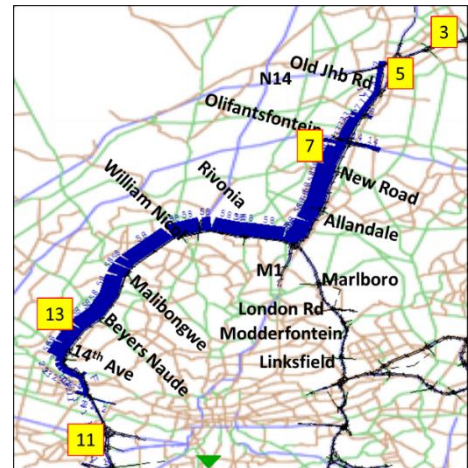


Figure 7: Trips between Gantry 13 and 7

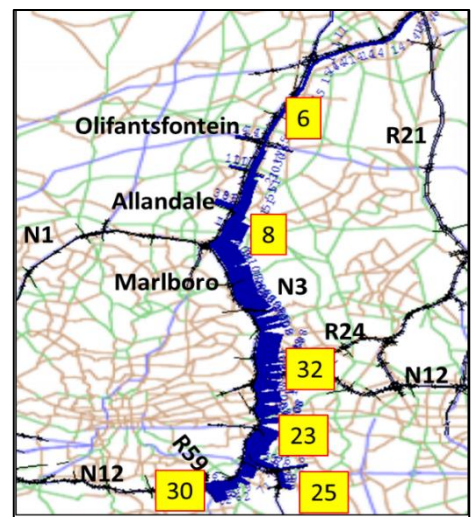


Figure 8: Trips between Gantry 8 and 23

Again, using the freeway cordon model, a select link analysis at Gantry 23 shows that there are trips from the north that route along the N1/N3 and the R21/R24/N3. This multi-path routing results in “leakage” past Gantry 8, and the trips picked up north of Gantry 8 in the previous example. These trips enter the system through Gantry 34 on the N12 to the east of the Gillooly’s Interchange. This assignment is shown in Figure 9.

By adding gantry 32 to the list of upstream and downstream gantries in the formula, the leak is “plugged” and only includes the trips between gantry 8 and 23. This result is provided in Figure 10.

Adding the potential leakage gantries to the sum of the external gantries eliminates the leakage. In this instance, adding Gantry 32 eliminates the trips that bypass Gantry 8 by using the R21. Whilst plugging the leak in this model is a relatively easy task, it may be substantially more intricate on the full model or if this model is to work in a city such as London, where there are more than 1500 cameras around the city network.

This result shows that there is a need to ensure that any potential leakage through alternative routes must be stopped by identifying the gantry location(s) between the selected start and end gantries (in this case 8 and 23).

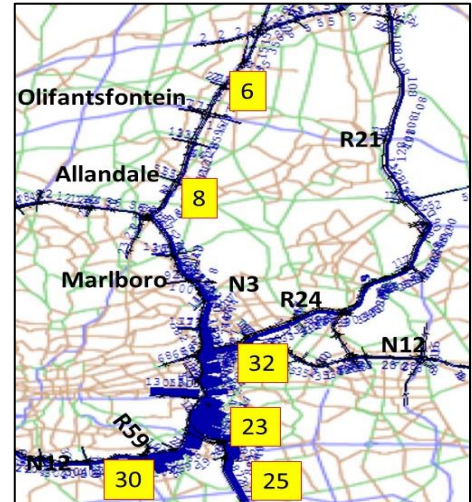


Figure 9: Select Link at Gantry 23

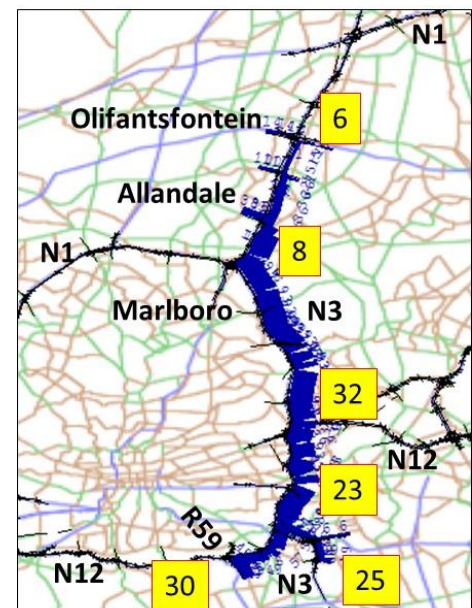


Figure 10: Trips between Gantry 8 and 23 only

5.4 ALTERNATIVE FORMULA FOR COMPLEX NETWORKS

The results above indicate that Equation 7 is applicable to relatively simple networks such as rural routes. Various alternative routes are used in an urban network especially when the primary routes are congested or when a stochastic assignment is used. Modifying Equation 7 eliminates the trips bypassing the selected gantries along alternative routes.

Equation 6 sums the trips in the select link matrices that pass through the gantries outside the route of interest. When these trips are subtracted from the common trips passing under the selected gantries, it zero’s the common cells associated with the selected gantries and produces

negative numbers in the un-associated cells. Retaining only positive numbers eliminates these trips from the resultant matrix.

In the example above where gantries 8 and 23 are selected, adding the select link matrix from gantry 32 to the “external” list of upstream and downstream gantries eliminated the leakage problem. Extrapolating this notion implies that excluding the select link matrices from all the external gantries would only increase the number of negative cells. Therefore, by only keeping the positive cell values would ensure the desired result.

There are two options to achieve this.

- List all gantries that are external to a desired route, sum them and insert this result in Equation 7 replacing the term $(\sum SL_{a-1} + \sum SL_{b+1})$
- Summing all select link matrices and subtracting all route-related select link matrices from this total as the replacement for Equation 7

The preferred option is a matter of the ease of computing all of the route combinations. The first option requires that lists of external gantries be prepared for each G2G pair. The second option only requires the definition of a route and all G2G combinations along the route are determined systematically.

This second option was selected for further analysis. Routes were selected along the freeway network, i.e. from Gantry 2 (N1 southbound north of the R21 Interchange) to Gantry 25 (N3 southbound R554 Heidelberg Rd Interchange).

Table 6 lists the G2G combinations derived from the route from gantry 2 to gantry 25. There are 36 G2G matrices derived from this route.

A batch file was created using the contents of Table 5 to call the necessary SL matrices and calculate the G2G matrix from the model.

Further work in the automation of this process is required to derive all probable G2G combinations through the network. This is achievable using the association table and the G2G Distance matrix. This will require additional software development, which is not considered within the scope of this work.

Table 6: G2G Combinations from Routes

G2G Matrix	From	To	Route								
			2	4	6	8	19	21	23	25	
1	2	2	2								
2	2	4	2	4							
3	2	6	2	4	6						
4	2	8	2	4	6	8					
5	2	19	2	4	6	8	19				
6	2	21	2	4	6	8	19	21			
7	2	23	2	4	6	8	19	21	23		
8	2	25	2	4	6	8	19	21	23	25	
9	4	4	4								
10	4	6	4	6							
11	4	8	4	6	8						
12	4	19	4	6	8	19					
13	4	21	4	6	8	19	21				
14	4	23	4	6	8	19	21	23			
15	4	25	4	6	8	19	21	23	25		
16	6	6	6								
17	6	8	6	8							
18	6	19	6	8	19						
19	6	21	6	8	19	21					
20	6	23	6	8	19	21	23				
21	6	25	6	8	19	21	23	25			
22	8	8	8								
23	8	19	8	19							
24	8	21	8	19	21						
25	8	23	8	19	21	23					
26	8	25	8	19	21	23	25				
27	19	19	19								
28	19	21	19	21							
29	19	23	19	21	23						
30	19	25	19	21	23	25					
31	21	21	21								
32	21	23	21	23							
33	21	25	21	23	25						
34	23	23	23								
35	23	25	23	25							
36	25	25	25								

The first step in this process is to calculate the sum of all select link matrices (SL_{ALL}). The second is to calculate the sum of the select link matrices along the selected route (SL_{Rt}). Equation 4 therefore becomes:

$$T_{External} = \sum (SL_{ALL} - SL_{Rt})$$

Equation 8

Substituting this into Equation 5 the G2G matrix based on a route definition becomes:

$$T_{G_{ab}} = \sum_{T_{ij}>0} \{\sqrt{SL_a \cdot SL_b} - (SL_{ALL} - SL_{Rt})\}$$

Equation 9

Where:

$T_{G_{ab}}$	=	Trips from Gantry(a) to Gantry(b) only
SL_a	=	Select link matrix from the Gantry(a) link
SL_b	=	Select link matrix from the Gantry(b) link
SL_{ALL}	=	Sum of select link matrix(ces) of all gantry locations
SL_{Rt}	=	Sum of select link matrices along the route from Gantry(a) to Gantry(b) inclusive

The example Gantry 8 to 23 was used as an example to compare the various methods. Initially, the number of trips in the model during the peak hour entering the network between gantry 6 and 8 and exiting the network between 23 and 25, was 107. Including gantry 32 on the N12 in order to exclude trips bypassing gantry 8 via the R21, resulted in 92 trips. The alternative formula above also produced 92 trips. The trips that only pass through gantry 1 was also tested and both methods resulted in 903 trips.

If trips between two gantries can chose two alternative routes, the ODs per route, or both routes, are derived by using:

- SL_{R1} = sum of SL matrices along route 1
- SL_{R2} = sum of SL matrices along route 2
- $SL_{R1,2}$ = sum of SL matrices along routes 1 and 2

5.5 ANPR DATA FOR IMPROVED TRANSPORT MODEL DEVELOPMENT

Comprehensive and accurate data is an important requirement for the development of traffic models. New technologies are replacing traditional manual data collection methods and the tracking of devices and/or vehicles through a network is relatively new. Any reliable traffic data is useful, provided the traffic model developer understands the usefulness and potential limitations of the data. The data is also only as good as the accuracy and reliability of any assumptions or factors applied to the data to make it useful for model input.

The ANPR data obtained from the GFIP ORT system is an extensive, reliable and accurate data source. Comparing the ANPR data to CTO data recorded on the same section of freeway as the ORT gantries shows that the ANPR data collects close to 100% of vehicles. Over 98% of the recorded vehicles have number plates that are traceable through the freeway network.

In recent studies, ANPR data is used to derive traffic counts, comprehensive journey times and OD information along “closed” road corridors with cameras located on cordon entry and/or exit points. The limited use of ANPR, Bluetooth tracking and traditional number plate surveys in an open system appears to stem from the knowledge that this data does not provide the trip’s actual

origin or destination. It is possible to identify cells, and the proportion of trips in these cells, that pass between specific locations (ANPR cameras). The ANPR data provides the traffic count between two locations and not further or through other monitored locations. The area within which the origins and destinations are situated is reduced significantly. Therefore, using the ANPR's counts over distance (G2G counts) to validate a model instils confidence in and, in fact, can be used to reinforce the trip distribution in the model.

Adding this count over distance data to the usable model database creates a significant arsenal of comprehensive, accurate and reliable data for transport model development. Currently, the use of traffic counts for trip matrix estimations has one significant potential downfall and can potentially distort the model's trip distribution. Using counts over distance would add accuracy to the trip distribution as opposed to detracting from it.

In fact, it can be argued that a model's initial trip distribution is synthesised from data sources with relatively small samples. Comparing commercially sourced probe data to traffic counts reveals a sample size of between 5% and 8%. Although the coverage is extensive, there is no reported information regarding the bias of the sample with respect to vehicles at the higher price range, i.e. potentially not including the proportion of road users that cannot afford this technology. Using the probe data does however provide a reasonable trip distribution starting point and with ANPR data to validate and potentially improve on the trip distribution, the OD trip matrices of transport models would be significantly more reliable and robust.

In summary, ANPR data has the potential to add significant reliability to the development of traffic models, including complex urban network models. This comprehensive data is essentially directly usable since it does not require expansion factors. In the case of the ORT system, the classification of the vehicles by type is an additional benefit and is therefore usable for modelling light and heavy vehicles.

The ORT system deployed in the GFIP project provides a continuous data stream that is usable for a post opening project evaluation (POPE) (Highways England, 2016) where projects are evaluated 1 and 5 years after opening. Such studies are sadly lacking in South Africa at this time. A POPE evaluation determines if the project objectives were achieved and in the case of a traffic modelling project it would be evaluated in terms of the accuracy of the forecasts. Finding significant errors should trigger investigations into the source of the errors to update the models through lessons learned and documented to benefit the traffic modelling fraternity.

A POPE exercise is provided in Section 6 below where the 2015 forecasts from the 2006 base year GFIP traffic model are compared.

6 CASE STUDY USING THE GFIP TRAFFIC MODEL

This section provides a brief overview of the GFIP project and the traffic modelling undertaken during the GFIP feasibility study. This model is used to provide 2015 forecasts and outputs that are directly comparable to the metrics derived directly from the ANPR data from the ORT systems (refer to Section 4.1). Comparing these outputs provides a means to validate the 2015 model traffic forecasts by:

- Comparing the model's traffic volumes with the gantry counts
- Comparing the model's journey times along the freeways with the measured G2G journey times
- Testing the formula for deriving G2G traffic count comparisons using the full GFIP model and comparing the model's distribution of traffic along the freeways with the G2G counts derived from the ANPR data

As this work involves the use of the GFIP traffic model's 2015 forecast, the process should be viewed as a proof of concept as opposed to expecting ANPR data to validate the model. The objective of using the GFIP model is to highlight elements in the model that require additional re-calibration when the model is next updated. At that time, the processes will be used to validate the new base year model and all time period and vehicle classes will be analysed.

6.1 TRAFFIC MODEL DESCRIPTION

In 2006, SANRAL appointed consultants⁴ to develop the GFIP traffic model to derive:

- A representation of current (2006) traffic conditions through the calibration and validation of the model
- Determine the attraction of traffic from the surrounding road network as a result of the capacity upgrades
- Develop traffic forecasts from land use forecast, economic forecasts and historic traffic data
- Develop an equitable toll strategy and an affordable toll tariff structure to ensure that the road user has a net benefit derived from savings in time and vehicle operating costs offered by the upgrades less the tolls charged

⁴ Goba (Pty) Ltd was the appointed consultant and Alan Robinson was the transport modeller that developed the GFIP traffic and toll revenue forecasting models

- Derive toll revenue forecasts over a 20-year period for input into the financial models.
- Provide design year traffic forecasts to the various design consultant's traffic engineers for the preparation of simulation models for the detailed designs of freeway sections and interchanges.

The study area for the GFIP model included all higher order (Class 1, 2 and 3) roads in the Gauteng province. At that time, the provincial GTS2000 traffic model was the latest version of the provincial transport model. This model was a multi-modal model with a focus on passenger transport and the most appropriate modelling tool was the EMME transportation modelling software. The GFIP model required a highway assignment focus and the preferred software selected was SATURN. After conversion of the model, further updates to the network included revising the freeway links to one-way pairs and updating the network's volume delay functions according to the limited survey data. This was an important step in the model update as the reliability of a toll model's route choice algorithms are crucial to ensure acceptable model calibration, validation and robust forecasts, as described in Section 2.1 above.

The 2005 and 2010 land use forecasts (based on the 2001 census) provided the basis for the model's initial trip generations and calculated by using the South African Trip Generation Rate Manual. The initial trip distribution functions were derived from telephone survey data and the previous GTS2000 trip distribution functions.

The calibration of the model comprised three steps:

- Starting with the GTS2000 matrix, the available traffic counts were used with repetitive matrix estimation procedures⁵ to produce an assignment with traffic volumes as close as possible to the measured counts. Adjusting the volume-delay function parameters to produce acceptable modelled journey times produced network with logical routing. Checking shortest routes, "forests" (routes from a zone to all other zones) and potential alternative routes were methods adopted to ensure a calibrated model network.
- The initial vehicle matrices were used as a gauge for the trip generation i.e. matrix totals. The initial assignments were compared to traffic counts, categorised by road class. Reducing the average trip length in the distribution model has the effect of "shifting" traffic from higher to lower order roads. Varying a global factor to the trip generations and attractions changed the quantum of trips on the network. It also has the effect of "shifting" traffic towards the higher order roads when modelling peak

⁵ Repetitive matrix estimation to traffic counts may produce the correct traffic volumes through the traffic count locations but will have a high probability of having a negative impact on the trip distribution in the trip matrix by potentially introducing a short trip distance bias.

periods due to the reduced congestion and higher speeds offered by the freeways. An iterative process, involving changing the average trip length and the global trip generation-rate factor, resulted in a reasonable correlation between the model assignments and traffic counts.

- Assessing outlier traffic counts to identify and fix any localised model network issues and finally using matrix estimation using approximately two thirds of the 2006 traffic counts, resulted in a calibrated base year traffic model.

The model was validated against the remainder of the traffic counts and re-checking the network routes and journey times. In addition, sensitivity tests highlighted the model's reactivity to changes in the input parameters, including the perceived value of time and vehicle operating costs.

6.2 DESIGN YEAR MODEL SCENARIOS

There were two future network scenarios and four design year scenarios. The networks comprised:

- The upgrading of the existing freeways and interchanges (GFIP Phase 1)
- Phase 2 expansion of the freeways to include new green-field routes

The design year matrices comprised:

- 2010, 2015, 2020 and 2025
- The matrices comprised six user classes (three light vehicles and three heavy vehicles) and five time periods

This study attempts to evaluate the scenario comprising the Phase 1: 2015 model network and the average AM peak hour traffic demands.

6.3 2015 GANTRY TO GANTRY OUTPUTS

In this section, the February 2015 ANPR data is used to derive and present the G2G output information. These outputs were then used as the comparative for the GFIP 2015 forecast model's results with the comparisons provided in Section 6.4 below. For ease of reference and due to the size of the full output tables, examples of the output are provided here. Appendix D contains the more comprehensive outputs.

6.3.1 Gantry Counts

The gantry counts relate to the vehicles that pass under each gantry. This data includes any vehicle with a missing or obscured number plate. The counts are extracted by location and vehicle class and in 15-minute time periods. The 15-minute counts are rolled-up into hourly time periods as shown in the example in Table 7.

Table 7: February 2015 Average Weekday Counts at Gantry 6, 7, 8 and 9

Gantry	6				7				8				9			
	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total
0	223	25	47	295	275	33	43	351	319	28	48	394	303	24	25	351
1	126	33	56	214	159	29	39	227	178	36	54	268	169	22	23	214
2	99	28	41	169	122	29	39	189	133	32	46	211	119	22	23	164
3	135	27	38	200	142	39	49	230	179	32	40	251	151	24	26	201
4	507	37	51	596	332	55	68	455	624	51	56	731	548	28	29	605
5	4617	88	78	4782	1347	90	121	1558	4515	141	94	4749	2841	59	41	2942
6	8135	119	64	8318	4353	170	107	4630	11448	221	89	11758	6925	79	35	7039
7	7904	136	79	8120	5531	190	86	5807	10923	257	95	11275	6959	87	37	7084
8	7153	210	94	7457	4809	288	146	5243	8726	373	117	9216	6331	165	62	6558
9	5693	266	113	6072	4972	416	175	5564	7635	453	148	8236	6062	266	72	6400
10	4818	262	124	5204	4941	403	154	5498	6702	428	158	7289	5382	277	72	5731
11	4817	274	126	5216	5013	360	147	5521	6562	434	157	7154	5113	293	74	5480
12	4852	294	133	5278	5285	341	145	5771	6691	462	160	7313	5146	302	74	5521
13	4600	289	135	5024	5570	343	144	6057	6700	469	171	7340	5185	296	76	5558
14	4406	303	135	4845	5968	333	139	6440	6548	480	162	7190	5156	314	74	5545
15	4875	310	128	5313	7909	310	129	8349	7709	462	155	8327	6229	321	77	6627
16	5192	277	120	5590	8460	249	103	8812	8494	386	143	9023	6422	301	68	6791
17	4791	209	113	5113	7822	243	114	8178	7378	307	137	7822	5452	242	66	5760
18	3240	127	107	3474	6278	174	112	6564	5100	186	120	5406	4266	141	56	4463
19	2143	87	86	2315	3536	102	90	3728	3309	120	94	3522	2868	82	45	2995
20	1479	64	81	1623	2326	78	87	2491	2256	86	84	2426	1905	55	40	2001
21	1127	55	74	1256	1696	60	76	1832	1662	69	79	1810	1477	42	37	1555
22	787	41	57	885	1205	52	69	1325	1168	53	61	1282	1182	36	34	1252
23	489	30	45	563	700	40	57	796	673	39	47	758	791	27	31	849
Total	82206	3591	2125	87922	88751	4425	2441	95616	115631	5605	2513	123750	86980	3507	1196	91683
Av 6-9	7731	155	79	7965	4897	216	113	5227	10366	284	100	10750	6738	110	45	6893

6.3.2 Gantry to Gantry Travel Times

The G2G travel times are calculated from the ANPR data records. The output is in seconds and then converted to minutes. Table 8 provides an example of the travel times between the first 20 gantries. Four minutes is entered on the diagonal denoting the assumed times a vehicle is on the freeway when only passing under one gantry. A refinement to this assumption could be to average the times between the upstream and downstream times, however this also may be erroneous and for the purposes of this work not necessary.

Table 8: February 2015 Weekday AM Peak Hour G2G Travel Times (minutes)

Times (min)		FROM: 7.00 to 8.00																			
To From	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1																					
2				7.9		21.0		32.1		38.4		49.0		56.8		60.1				37.5	
3	7.5																				
4						14.7		26.1		32.3		45.0		50.6		55.3				30.9	
5	13.3		5.8																		
6								11.9		18.8		29.6		36.9		41.0				17.4	
7	18.6		10.8		5.2																
8										6.1		16.4		23.5		27.2				5.1	
9	26.3		18.0		12.4		7.7													4.5	
10												9.8		16.3		20.6					
11	44.3		35.3		30.0		24.8		17.2											21.7	
12														6.6		10.7					
13	58.2		50.1		43.5		38.7		32.1		15.3									37.4	
14																4.3				40.4	
15	64.2		57.7		52.0		46.6		39.6		23.2		8.8							46.3	
16																				40.0	
17	74.5		68.1		62.1		57.1		48.8		33.4		17.2		9.6			46.8	53.5	42.3	
18	28.2		19.8		14.0		8.7			5.3		15.5		23.2		25.4					
19																					
20	33.6		25.0		19.2		14.0			11.0		21.4		28.1		39.0		5.6			

Consecutive times between gantries can be used to validate the modelled journey times along the various sections of the freeways.

6.3.3 Gantry to Gantry Travel Speeds

Average speeds are calculated by dividing the distances by the times. Table 9 provides a sample of the average speeds of all vehicles on weekdays between 07:00 and 08:00. The matrix diagonal cells contain the average speed calculated from the remaining cells. An alternative method is to use the speeds from the sections leading to and from each gantry.

Speed is a recognisable metric to allow the modeller to appreciate the operations of the network. The results indicate that where the traffic flows are high, the resulting congestion and low speeds. Where flows are lower, in the counter-peak direction, the recorded speeds are noticeably higher.

The average G2G speed is comparable to the model network combined link speeds encompassing all links between gantry locations.

Table 9: February 2015 Average weekday speeds between 07:00 and 08:00 (kph)

To From	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																				
2				75.8		53.1		55.8		58.2		59.1		62.8		66.6				58.8
3	92.1																			
4						35.1		45.7		50.7		51.0		58.8		61.5				52.0
5	95.2		100.4																	
6								56.8		59.6		60.1		66.6		70.4				62.9
7	96.2		101.8		101.4															
8										72.0		67.3		75.7		81.3				81.6
9	93.8		98.7		97.0		88.3													102.3
10												67.7		82.1		86.1				
11	69.5		67.6		60.4		51.8		35.5											49.5
12														102.5		103.6				
13	62.7		59.0		54.6		47.9		36.7		37.1									43.8
14																100.4				62.5
15	64.7		60.1		55.6		50.8		42.6		46.6		58.6							48.6
16																				52.4
17	62.7		58.4		54.8		50.4		45.1		47.7		59.8		53.3			52.7	49.8	49.1
18	88.9		91.9		88.7		81.9			84.4		71.2		76.8		87.4				
19																				
20	86.2		88.3		84.9		78.9			76.3		70.0		77.5		66.9		69.9		

6.3.4 Gantry to Gantry Counts

The extraction of G2G counts for each hour in the day, for any combination of days and any or all vehicle classes is possible using the Data Miner software. These counts relate to the distances between the matrix combinations and are therefore counts over distance.

Table 10: Gantry to Gantry Counts for All Vehicles on Weekdays - 07:00 to 08:00

All Vehicle Counts FROM: 7.00 to 8.00																				
To From	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1784																			
2		4157		1566		586		451		63		26		7		13				26
3	986		3124																	
4				831		199		137		20		8		3		3				14
5	263		700		1626															
6						3123		2397		321		140		43		51				265
7	276		611		1425		1023													
8								3461		872		363		85		59				577
9	62		166		279		233		1706											664
10										791		426		83		56				
11	27		55		115		140		884		2408									281
12												778		382		384				
13	6		6		20		29		191		560		1698							48
14														977		1176				
15	3		5		19		14		51		244		1006		1541					1
16																740				1
17	6		4		16		10		35		126		654		1087		1998	1	1	11
18	23		40		98		112			332		114		13		2		499		
19																				513
20	27		35		130		190			604		178		8				1259		1074

There is no direct relationship between these counts and standard model outputs. However, the formulae described in Section 5.1, makes it possible to compare these counts over distance and equivalent values derived from the traffic models.

6.3.5 Trip length frequency distributions

Combining the counts over distance and the distances between the gantry pairs provides a trip length distribution. Figure 11: G2G Trip Length Distribution depicts the G2G trip length frequency distribution. Because the G2G trips are not full trips from the trip origin to the destination, the direct comparison to standard model output is not possible. The assumed diagonal distance also may distort this result.

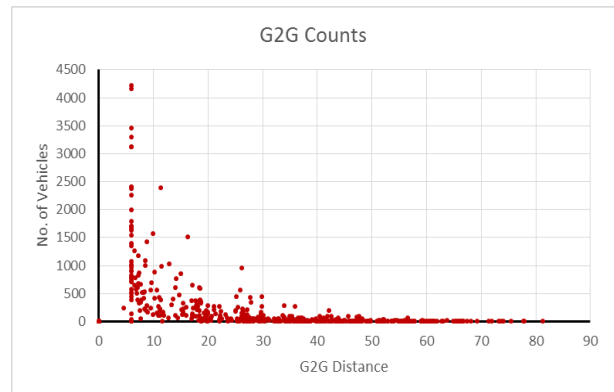


Figure 11: G2G Trip Length Distribution

Figure 12 displays the trip length distribution without the diagonals.

These results highlight the high reliance on the freeway network for short distance trips. Note that these results also exclude those trips that use the freeways but do not pass under any gantries.

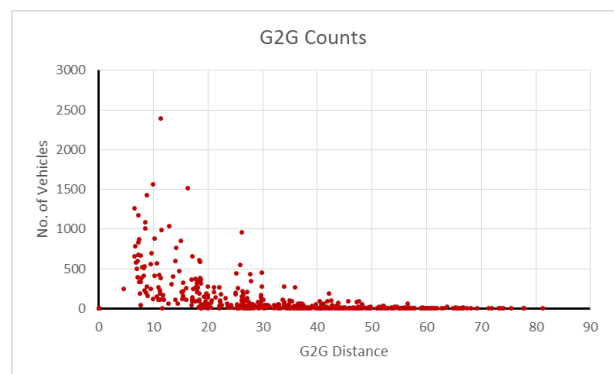


Figure 12: G2G Trip Distribution without Diagonals

6.4 COMPARISON TO MODEL OUTPUTS AND MODEL IMPROVEMENT OPTIONS

The comparison between the GFIP traffic model's outputs and the ANPR data provides a basis to either validate the 2015 forecast model as a new base year model or a measure of the model's forecasting robustness in a Post Opening Project Evaluation. The comparisons should also provide valuable insight into potential model improvements implementable during an update of the model.

The comparisons provided below are between the February 2015 AM peak hour ANPR data and the 2015 GFIP model forecasts and compares actual data against a modelled forecast that is nine years after the development of the 2006 base year model. Although the design year infrastructure (10 and 12-lane highways) had been modelled, no traffic count or journey time data existed in South Africa for developing volume delay functions. Furthermore, the capacity improvements to the Gauteng freeways has promoted significant land use development at key development

nodes, which had not been foreseen at the time of developing the forecasts. Therefore, results from the validation processes performed on the 2015 model forecasts are only intended to demonstrate the validation processes using the ANPR data and identify improvements that could be made to the GFIP when it is next updated and re-calibrated.

The comparisons are made using the Transport Analysis Guidance (webTAG) (Department for Transport, 2014) validation criteria, which are commonly used in South Africa.

6.4.1 Traffic Volumes

Table 11 and Figure 13 provide the comparison between the average traffic volumes for the AM peak period (0:600 to 09:00) from the actual gantry counts (February 2015) and the GFIP traffic model.

Summing all the counts for comparative purposes indicates that the traffic model forecast is approximately 9% too low for light vehicles and 22% too low for heavy vehicles. Scrutiny of the light vehicle traffic volumes shows that there are variations in the percentage difference between the actual and modelled flows. However, an average forecast of 10% under actual traffic for a nine-year forecast horizon on an urban open road toll system is certainly within acceptable limits. International averages for similar projects have resulted in actual traffic volumes being 20% to 25% below forecasts (Standard & Poor's, 2005).

Possible reasons for the higher than forecast outcome could be a result of:

- Social defiance of the open road toll system. Whilst many road users are not meeting their legal obligation to pay, the decision not to pay could have a cost or value of time associated to it. If enforcement ensures that all road users pay, what would be the toll-road attraction rate be? Initially, one could expect a significant reduction but, with the congestion on the alternative road network, the ramp-up period could be short.
- The model input includes estimated values of time from surveys on rural roads. These may have been slightly low when compared to users on urban freeways.
- Modelled speeds on the highways being too high in comparison to the relative speeds on the alternative routes making the tolled highways more attractive.

Table 11: Comparison of Gantry and Model Traffic Volumes (Average 06:00 to 09:00)

Location	Gantry	Actual			2015 Model Forecast			Lt %diff	Hgv %diff	Total %diff
		Light	Small HGV	Large HGV	Light	Small HGV	Large HGV			
Lynwood - Proefplaas	1	3728	176	146	3212	104	106	-14%	-35%	-15%
Rigel - Flying Saucer	2	7992	175	116	7595	132	93	-5%	-23%	-6%
Botha - Flying Saucer	3	5665	148	79	4841	118	55	-15%	-24%	-15%
John Vorster - Brakfontein	4	4795	119	62	4472	81	40	-7%	-33%	-8%
Rooihuiskraal - Brakfontein	5	5244	194	91	4443	171	80	-15%	-12%	-15%
Samrand - Olifantsfontein	6	7731	155	79	7221	166	64	-7%	-2%	-6%
New Road - Olifantsfontein	7	4897	216	113	4105	133	53	-16%	-44%	-18%
Allandale - Buccleuch	8	10366	284	100	8584	209	88	-17%	-23%	-17%
Rivonia - Buccleuch	9	6738	110	45	6352	82	32	-6%	-26%	-6%
Rivonia - William Nicol	10	4841	195	57	4440	127	37	-8%	-35%	-10%
Hans Strijdom - William Nicol	11	6157	105	46	6324	116	60	3%	16%	3%
Hans Strijdom - Beyers Naude	12	3028	132	46	2718	85	37	-10%	-32%	-11%
14th Avenue - Beyers Naude	13	5977	148	68	4915	117	68	-18%	-14%	-18%
Gordon - Maraisburg	14	3871	124	55	4367	115	48	13%	-9%	12%
Soweto Highway - Maraisburg	15	6327	167	87	5452	136	82	-14%	-14%	-14%
Randshow - Old Potch	16	3531	116	52	3043	75	32	-14%	-37%	-15%
Golden Highway - Old Potch	17	4700	108	93	3824	118	116	-19%	17%	-17%
Marlboro - Buccleuch	18	5776	257	109	5422	176	94	-6%	-26%	-7%
London - Modderfontein	19	5785	179	69	4928	135	49	-15%	-26%	-15%
Modderfontein - London	20	7301	243	119	6779	252	122	-7%	3%	-7%
M16 Linksfield - N12 East	21	6161	223	87	5962	148	58	-3%	-33%	-5%
van Buuren - Gilloolys	22	6630	312	296	6008	286	212	-9%	-18%	-10%
M2 Geldenhuis - N3	23	3986	287	185	3665	205	140	-8%	-27%	-10%
Rand Airport - Elands	24	4466	174	213	3723	109	169	-17%	-28%	-18%
M94 Grey - Nederveen Highway	25	1411	103	85	1487	71	66	5%	-27%	2%
De Villiers Graaff - N1	28	4124	221	121	3485	105	66	-15%	-50%	-18%
klipriver - comaro	29	3743	103	72	3262	67	55	-13%	-30%	-14%
Reading - Comaro	30	3113	183	108	2586	116	79	-17%	-33%	-18%
Sybrand Van Niekerk Freeway - Voortrekke	31	4918	131	140	3856	102	78	-22%	-34%	-22%
R24 / Boeing Road - N3	32	7143	343	191	7805	367	169	9%	0%	9%
M44 Jet Park Road - R21	33	2371	173	88	1851	133	92	-22%	-14%	-21%
Rondebult Road - Atlas Road	34	5711	139	101	6107	116	47	7%	-32%	5%
atlas - Tom Jones	35	1810	146	78	1502	90	75	-17%	-26%	-18%
Hans Stijdon - N1	37	1758	50	32	1779	33	20	1%	-36%	0%
M31 Nelmapius - N1	38	3860	215	117	2690	161	100	-30%	-21%	-30%
Olifantsfontein - M31 Nelmapius	39	3249	214	114	2947	197	122	-9%	-3%	-9%
Irene - Olifantsfontein	40	3738	83	86	3382	60	68	-10%	-24%	-10%
R25 - R23 Benoni	41	4713	145	114	4730	112	106	0%	-15%	0%
R25 - Olifantsfontein	42	3698	259	159	3774	231	151	2%	-9%	1%
Griffiths - OR Tambo	43	2588	181	74	2159	117	56	-17%	-32%	-18%
Griffiths - N12	44	2293	97	50	1792	54	30	-22%	-42%	-23%
24 - M96 Voortrekker	45	1713	150	99	1930	120	77	13%	-21%	8%
Totals		197646	7283	4243	179521	5650	3392	-9%	-22%	-10%



Figure 13: Comparison of traffic volumes by vehicle type and location

At first glance, the forecasts appear reasonable and for a tolled traffic forecast, an average of -10% lower than actual would be an acceptable result. However, the current GFIP model will require further calibration going forward to become the validated base year model. Table 12 provides the results of the model versus ANPR data using the webTAG criteria. The comparison measured and modelled traffic volumes using the GEH statistic (Equation 10). A GEH of 5 or less for 85% of observations is considered acceptable. However, on large models where there is limited time and budget constraints a GEH of 10 or less for 85% of the observations is sometimes used, but is not ideal.

$$GEH = \sqrt{(V_2 - V_1)^2 / (0.5(V_1 + V_2))} \quad \text{Equation 10}$$

Table 12: Validation Criteria Results

Flows 700-2700vph						
Criteria	<15%		<20%		<25%	
Count	OK	%	OK	%	OK	%
7	3	43%	5	71%	7	100%
Flows > 2700vph						
Criteria	<400		<650		<900	
Count	OK	%	OK	%	OK	%
35	13	37%	23	66%	31	89%
GEH						
Criteria	5		10		15	
Count	OK	%	OK	%	OK	%
42	10	24%	28	67%	39	93%

In terms of the webTAG criteria, the model does require further calibration, even if one was to accept the GEH 10 for 85% of observations. Using less stringent criteria values provides an indication as to how far out the model is. Whilst it does not meet the validation criteria, it will be a suitable prior model for the model update.

6.4.2 Journey times and speeds

Calculating the journey times between gantries using the ANPR data is accurate whereas the modelled journey times are measured from the closest nodes to the gantries. Gantries are located mid-link, which do not necessarily coincide with the node locations. The longer distance between measured gantry locations, lower the potential error introduced by distance between the gantry and the closest node.

Comparisons between the following data sets from the ANPR data and the GFIP model forecasts include:

- The average speeds for light vehicles that start their trips between 07:00 and 08:00.
- The average speeds for heavy vehicles that start their trips between 07:00 and 08:00.

Figure 14 depicts the comparison of the light vehicle ANPR measurements with a 15% variation (black line) and the modelled journey times (red line). The routes include the N1 in the northbound direction between the golden Highway Interchange (after Gantry 17) and the Proefplaas Interchange (after Gantry 1), and southbound between the R21 Flying Saucer Interchange (after Gantry 2) and the Rand Show Interchange (after Gantry 16).

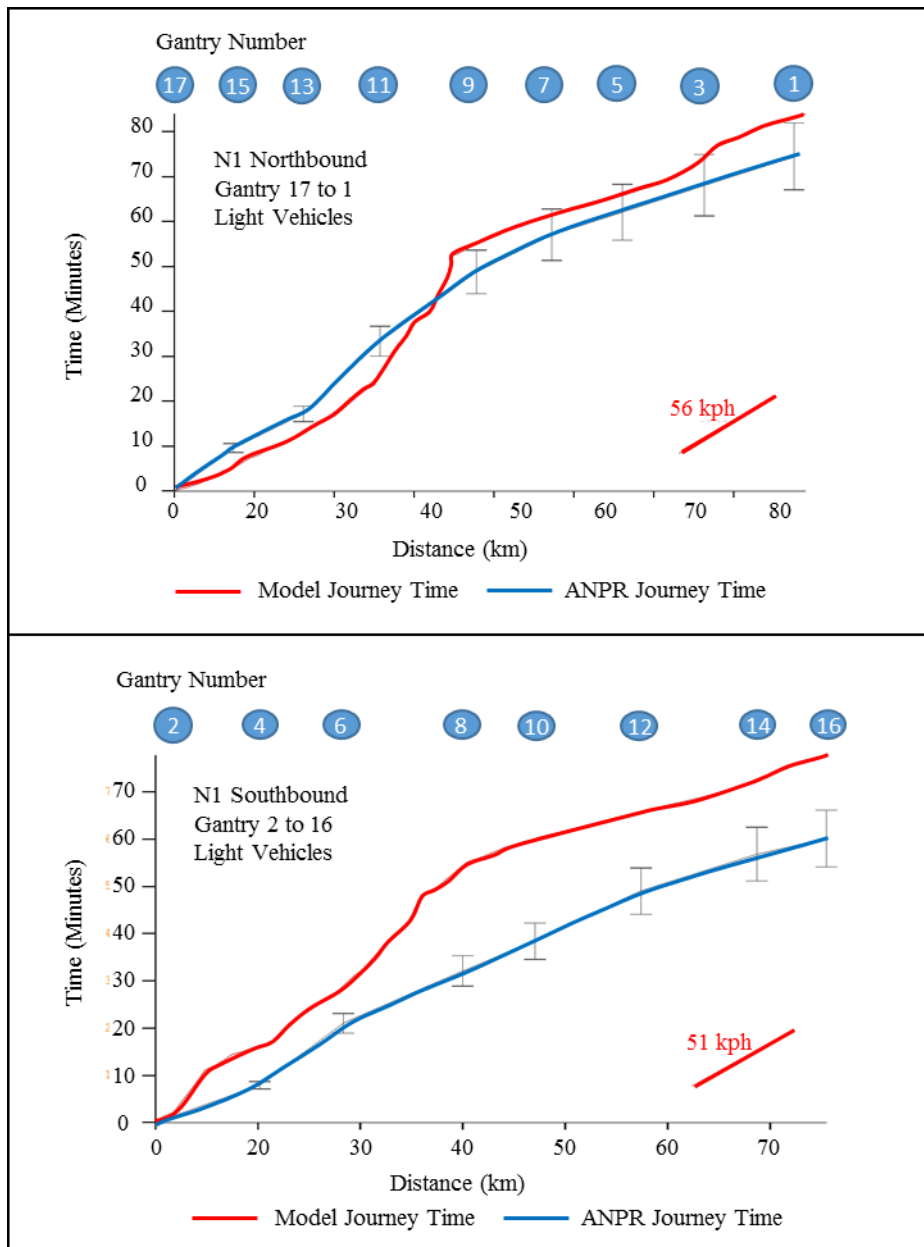


Figure 14: Model versus Gantry Journey Times on the N1

These results highlight the following regarding the modelling of the journey times, or inversely the modelled operating speeds:

- For many sections of the freeways the travel speeds are within acceptable limits and shown where the measure and modelled speeds are the same, i.e. the lines are parallel.
- An over-estimation of modelled delays because of congestion. This is highlighted by the steep increase in times between Gantry 2 and 8 in the southbound direction and Gantry 13 and 9 in the northbound direction

The results for the N3 route and those relating to heavy vehicles are provided in Appendix H

Based on these results, the improvements to the traffic model would include:

- Checking the network for the correct cross-section (number of lanes) where there are large discrepancies
- Checking the errors in the traffic volumes as relatively small changes in volumes can significantly increase delays when operating at congested conditions
- Increasing the link-type capacity to delay the onset of the increase rates of change in delays
- Lower the BPR power values, which smooth the transition between free-flow and at capacity / over-capacity conditions.

6.4.3 Gantry to gantry (G2G) counts

Comparing the G2G counts and the equivalent values calculated from the 2015 model forecasts shows the correlation between the data sets for light and heavy vehicles. The average slope of the graph for light vehicles is 0.94 indicating that the quantum of trips in the model is marginally lower than the ANPR data whereas for heavy vehicles the slope is 0.72 indicating that the modelled values are approximately 28% lower than the ANPR data. Figure 15 and Figure 16 show these results.

The comparison on a G2G cell level is provided in matrix format in Appendix H.

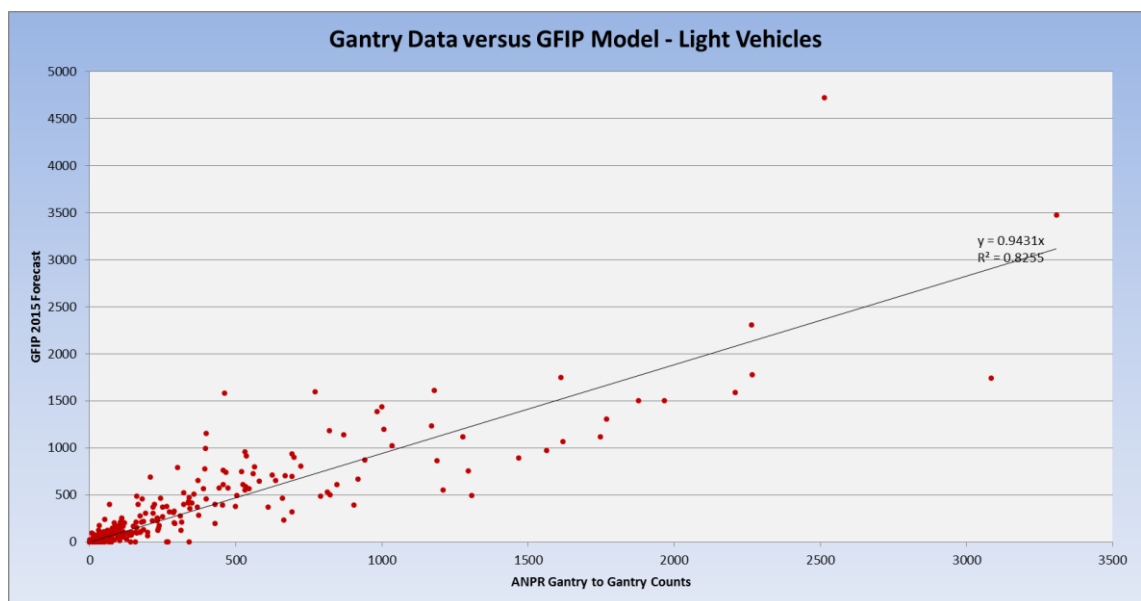


Figure 15: G2G vs Model Output - Light Vehicles

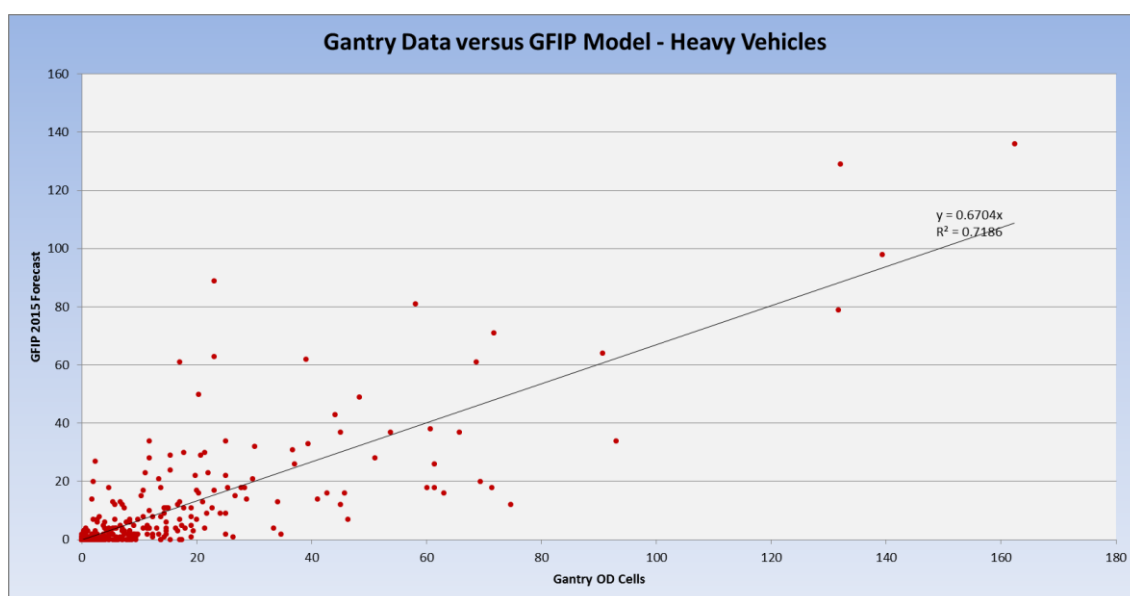


Figure 16: G2G vs Model Output - Heavy Vehicles

Whilst there are no documented acceptable criteria for count over distance, measured and modelled traffic volumes GEH statistic (Equation 10) has been used here. Table 13 below summarises the results of the comparison between the G2G counts and the equivalent values from the GFIP model's 2015 forecast for the AM peak period.

Table 13: Summary Statistics for G2G and Model Count Comparison

Vehicle Class	Slope	R ²	GEH <5	GEH <10
Light	0.9431	0.8255	67%	88%
Heavy	0.6704	0.7186	91%	100%

Whilst the light vehicle comparison yields a closer correlation in terms of the R² and a slope close to one, the GEH statistic indicates that the improvement in the distribution of trips through the freeway network will improve the reliability of the model.

The heavy vehicle results confirm that too few heavy vehicles are assigned to the freeways. The GEH statistics show an “acceptable” correlation, but this is only because of the low values in the heavy vehicle traffic. The R² results indicate significant variations between the ANPR counts and the model. Note that the heavy vehicle distribution was based on a previously untried method of using land use data relating to industrial and retail land uses. This was, at the time, deemed an improvement on the constant percentage of the trip matrix used in previous models.

Adjusting the heavy vehicle trip matrices according to G2G counts would significantly improve the modelling of heavy vehicles.

6.4.4 Trip Length Frequency Distribution

Figure 17 depicts the trip length frequency distribution on the freeways for light and heavy vehicles during the morning peak period. Both vehicle classes show that the modelled values are lower than the ANPR data indicating that the average trip length on the freeways is longer than the model predicts. This is more evident in the heavy vehicles. The normalised distributions relating the percentage of trips and the distance travelled shows a less prominent difference.

The average trip length is an important element in the trip distribution and the comparison of the modelled and ANPR. Table 14 provides the comparison of the average trip lengths for light and heavy vehicles. Note that in this calculation, an average trip length of 6km was used for the measured and modelled short distance trips that only pass one gantry location (refer to Section 4.3)

Table 14: Measured and Modelled Average Trip Length on the GFIP Network

Vehicle Type	ANPR Data	Model
Light	11.30 km	11.43 km
Heavy	13.11 km	12.11 km

Note: The estimated average trip length for a trip that passes one gantry is 5.5km.

The average light vehicle trip length travelled on the GFIP network compared well with that measured using the ANPR data. This would be because of the use of the telephone survey data obtained at the commencement of the modelling work. The heavy vehicle's average trip length was an unknown at the time of developing the traffic model and assumed to be longer than that of light vehicles. Whilst this is the case, the assumption turns out to be an underestimation and an improvement opportunity for the next model update.

The percentage of trips in the ANPR and model that travel various distances was compared and the results are provided in Table 15. These results show that the model represents the distribution of trips on the freeways reasonably well.

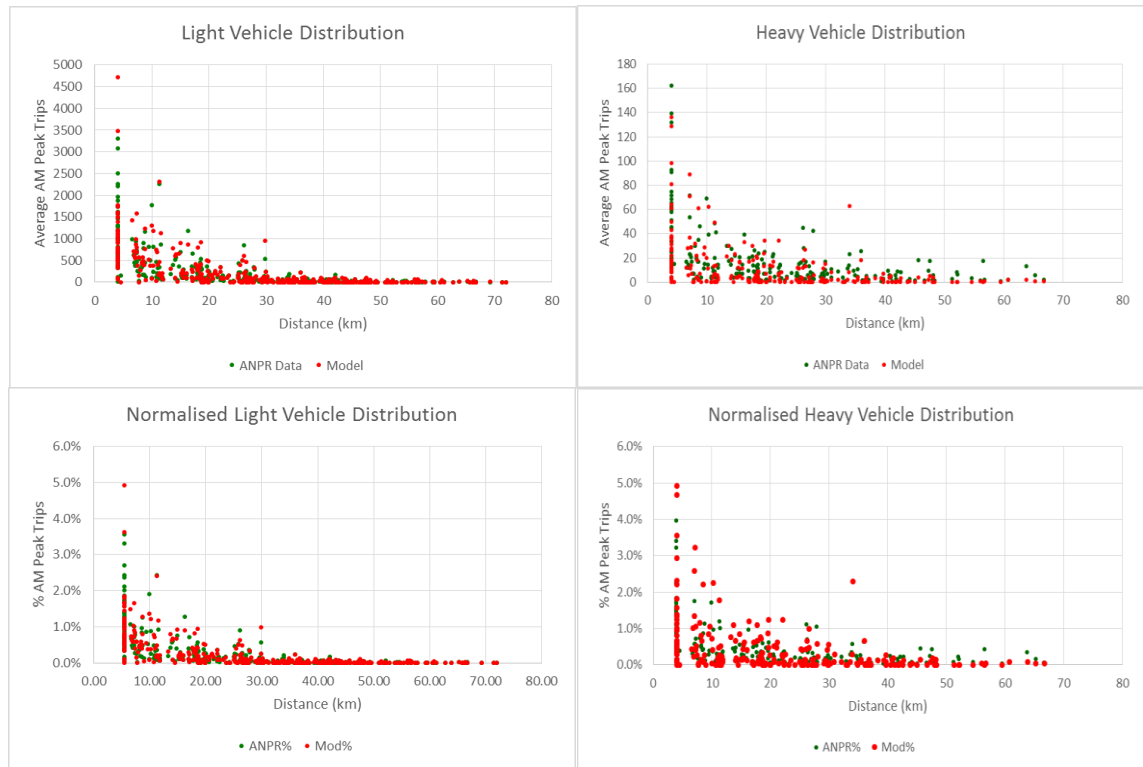


Figure 17: ANPR and Model Trip Length Frequency

Table 15: Comparison of percentage of trips versus distance travelled

Distance Travelled	ANPR Data	Model
<10km	65.8%	62.9%
>10 km and < 20km	20.3%	23.8%
>20 km and < 30km	9.4%	9.1%
>30 km and < 40km	2.8%	2.5%
>40 km and < 50km	1.3%	1.2%
>50km	0.5%	0.5%

6.5 POTENTIAL GFIP MODEL IMPROVEMENTS

Updating the GFIP traffic model to provide traffic forecasts for the next phase of the Gauteng Freeway Improvement Project will benefit from the insights that the analysis of the ANPR data provided. Identifying potential improvements to the current GFIP traffic model was one of the objectives of this study.

Using the comparison of the model outputs and ANPR data, improvements to the original GFIP traffic model are possible in the following areas:

- Improving the modelled road network based on the journey time data, especially on parts of the freeways impacted by gradients and/or high weaving that reduces capacity.
- Improving the trip generation rates applicable to the strategic model with its large zones, especially in relation to heavy vehicle movements.
- Improving the trip distribution using the average trip length from the sample data in the derivation of the gravity model and the trip length frequency distribution

From a network perspective, the journey time information through the freeway network provides valuable insight into locations where the actual and modelled speeds differ. Whilst the modelled times are longer than the measured times, the data does not give a definitive reason for the difference. Potential reasons could be:

- Table 11 indicates that the modelled volumes on the freeways are too low, especially regarding heavy vehicles. Possible reasons include:
 - The capacity restraint on the secondary road network requires relaxation. Journey time information pertaining to the alternative routes is required
 - Trip generation rates are too low. Using screen-line counts that include all routes serving key corridors would verify this.
 - Average trip length too short, resulting in a bias towards the lower order roads
- Figure 14 indicates that the modelled network is operating too slowly under the modelled traffic volumes. Since it was established that the traffic counts at the gantries are higher than the assigned volumes, the following needs further investigation:
 - Traffic counts on freeway links between gantries to check if these links have additional short distance trips assigned to them.
 - Select link analysis from links between gantries to check for additional short-distance trips.
 - The operating speeds of alternative routes that run parallel to the freeway in the corridor adjacent to problematic freeway sections. If there is insufficient capacity in the modelled alternative routes, possibly resulting in the additional short-distance freeway trips.
- Figure 15 and Figure 16 show that there are some noteworthy outliers in the comparison of the G2G data from the ANPR and modelled outputs. The results shown in Figure 17 supports the fact that there are too many short distance trips in the model. This will require an adjustment to the model's distribution function. Whilst the inputs into the original model may have been applicable at the time of the model development, the

additional capacity of the upgraded freeways may have resulted in an increase in the average trip length, which is measured in the ANPR data. In most modelling works, the base year trip distribution, i.e. average trip length (calibrated into the constant β from survey data) is assumed constant. These results provide the necessary evidence to reassess this assumption.

It is very important that the speed flow relationship, as defined in the volumes delay or BPR (Bureau of Public Roads; US) curves is as accurate as possible as this has an impact on trip routing, particularly under congested conditions. Incorrect trip routing will introduce errors in the SL matrices. These errors can be exacerbated using standard matrix estimation to traffic count techniques.

The ANPR data obtained from the ORT systems on the Gauteng freeways has been used to highlight deficiencies in the current version of the GFIP traffic model. However, in terms of undertaking further investigation to find remedies, the ANPR data does not provide all the answers. Additional data from electronic count stations between the toll gantries on the freeways and on alternative routes is needed to confirm the problem and provide the data on which to base a solution. Journey times on the alternative routes from probe data is required to check the operational characteristics of the modelled alternative routes.

Aggregating the results of comparisons can smooth the averages and, in general, seem acceptable. Using the G2G counts in a similar way to examining problematic traffic count comparisons, potential improvements to specific areas or routes across the modelled network can be highlight. It is not possible to identify these modelled problem areas using any other data set currently available.

Table 16 highlights the worst GEH results from the G2G comparisons and they reflect a randomness in the outcome from the standard four-step modelling process. To illustrate this point these results show the following:

- Gantry 32 is on the westbound approach to the Gillooly's Interchange on the N12. Of the 7 143 trips passing under this gantry (Table 11), 2 513 do not pass under any other gantry, i.e. are short distance trips. The model overestimates this number of short distance by approximately 2 200 trips.
- Gantry 8 is on the southbound approach to the Buccleuch Interchange on the N1. There are approximately 10 300 peak hour trips (Table 11) of which 3 804 are short distance trips not passing under any other gantry. The model predicts approximately half this at 1 742 trips.

- Gantry 3 to 41 would pick up trips between the Centurion and Kempton Park areas using the R21. The ANPR data recorded 341 trips during the morning peak period, yet the model does not reflect any of them. This could be a model routing problem or a trip distribution problem.

Table 16: Worst GEH Results from G2G Comparisons

Gantry From	Gantry To	ANPR	Model	Diff.	GEH
32	32	2513	4719	-2206	36.69
19	21	462	1583	-1121	35.06
8	8	3084	1742	1342	27.32
19	19	397	1157	-760	27.25
31	31	1307	491	816	27.21
3	41	341	0	341	26.12
14	14	772	1598	-826	23.98
34	18	267	0	267	23.12
34	20	264	0	264	22.96
17	29	208	692	-484	22.80
32	20	397	998	-601	22.76
13	13	1208	555	653	22.00
19	33	67	401	-334	21.86
15	11	302	791	-489	20.94
17	13	664	236	428	20.19
17	15	903	396	507	19.90
4	8	160	484	-324	18.08
33	43	156	3	153	17.14
28	28	1296	758	538	16.79
34	10	139	0	139	16.69
11	11	1748	1118	630	16.63
22	22	1563	972	591	16.61

The comparison of the model output to standard traffic count and trip distribution data may not indicate that the model performs to an acceptable degree of accuracy. From the traffic counts, the total volumes passing under each gantry is acceptable. From a trip distribution perspective, the average distribution is within an acceptable degree of accuracy when compared to the initial sample data. However, the analysis of each of the worst ANPR G2G data comparisons illustrates potential problems in the model, which are “averaged out” in these standard calibration tests. The ability to disaggregate a traffic count into counts over different distances adds significant value to the checking of a model’s traffic distribution and/or trip routing.

The methodology to extract equivalent assigned volumes over the same distances, as described in Section 5, will add a powerful validation tool to the modelling process. The process enables the identification of the trip matrix cells, and portions thereof, that contribute to the count over a distance equivalent to the positions of the ANPR cameras. Knowing the actual count that corresponds to the total number of trips in this extracted matrix, the derivation of a factor to adjust this sub-matrix to the count is possible. Therefore, if the count over distance submatrix derived from the model is subtracted from the overall trip matrix, factored to the actual count over distance and added back into the model matrix, this would be an improvement in this one count value. Whilst acknowledging that this would improve one count over one distance, it could adversely affect others.

It is however similar to using one link count in a matrix estimation routine. Trips passing the single count location, represented in one select link matrix, is factored so that the matrix total matches the count.

The following figures illustrate the difference between matrix estimation to a link count and G2G counts using the link where gantry 13 is located, being northbound on the N1 between the 14th Avenue and Beyers Naude Interchange

Figure 18 illustrates the location and quantum of trips generated from the blue origin zones (upstream of the count) and the destination zones (downstream of the count in red) that may be included in the factored matrix. The volumes represent the average three-hour AM peak (0:600 to 09:00) for the 2015 gantry count and the 2015 forecast from the model. Correcting the model to match the count require factoring all ODs by 1.21 irrespective of the location of the origins and destinations.

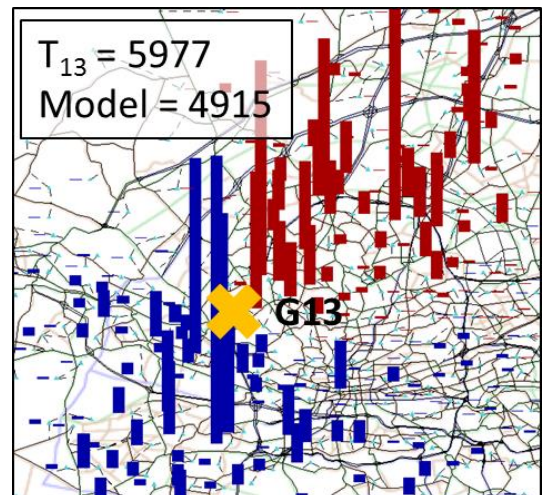


Figure 18: Select Link Matrix ODs through Gantry 13

Using ANPR data and G2G counts over distance as an alternative isolates smaller areas or pockets of origins and destinations across different parts of the network that make up the G2G count. Figure 19 illustrates the ODs that are associated with specific G2G combinations from Gantry 13.

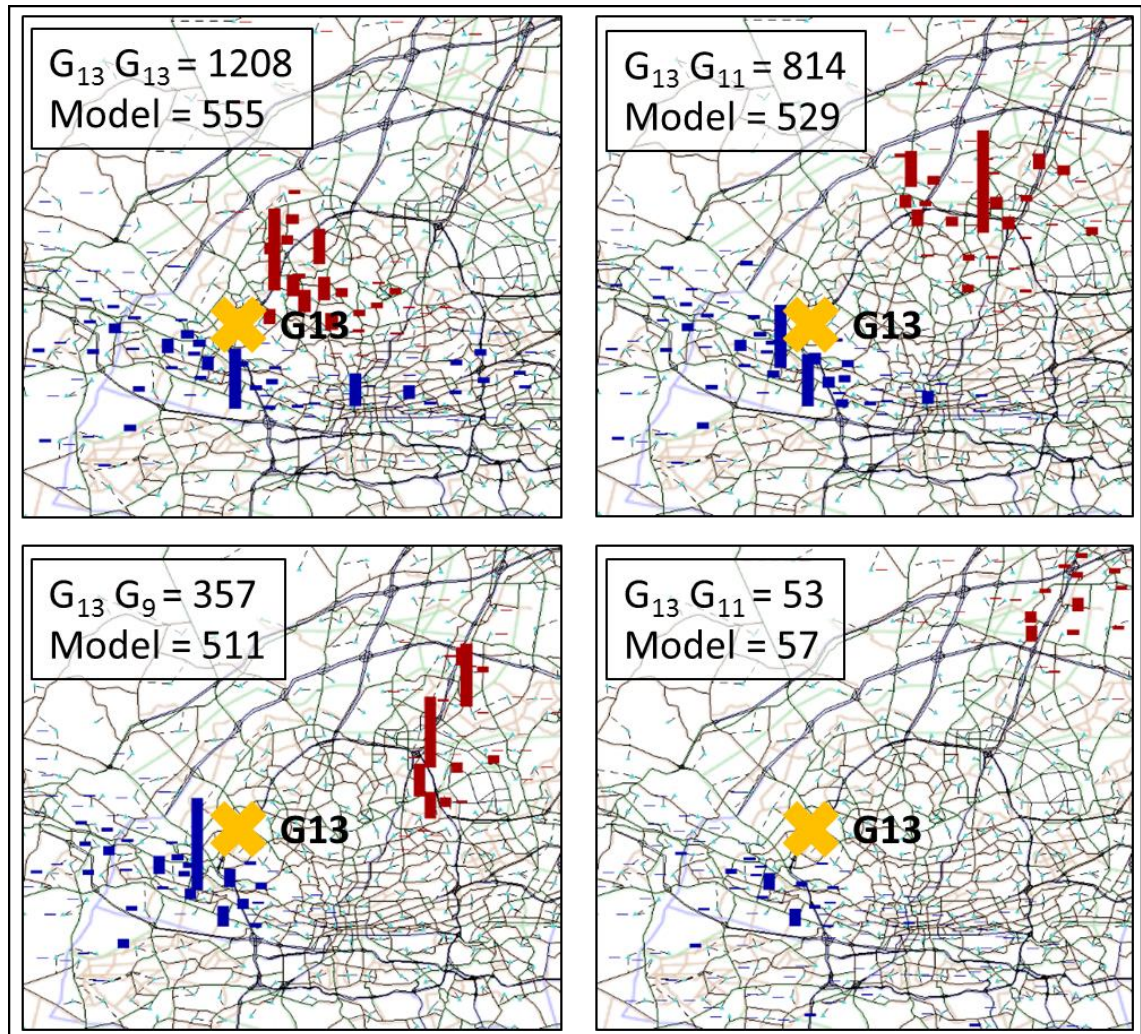


Figure 19: G2G Matrices from the GFIP Model and Actual Values

Each plot provides the actual G2G and the modelled traffic volumes. Based on these numbers, different factors would apply to each G2G combinations ranging from 0.7 to 2.1. The overall adjustment of the G2G matrices associated with Gantry 13 would have the net effect increasing the select link matrix factor of 1.21 to match the Gantry 13 count. The different factors for different groupings of OD pairs distributed throughout the network will result in a considerably better result by improving the model's trip distribution in addition to correcting the quantum of trips below each gantry.

The GFIP traffic model can be improved by adopting the above process to adjust all G2G model/count comparisons in the following manner:

- Subtract a SL2SL (G2G) matrix for a specific user class from the overall trip matrix.
- Derive the factor from the matrix total and the G2G count

- Factor the SL2SL matrix by this factor and add the result back into the overall matrix
- Repeat the above for all G2G counts.

On completion of the above, the assignment of the resultant matrices will be checked against other independent counts. If successful, this process will be included in the development of the 2015 base-year trip matrices and further model validations. However, undertaking this work is beyond the scope of this study and will be work undertaken by the model owner.

It is possible for the incorporation of this method into current matrix estimation techniques (Ortúzar & Willumsen, 1998) and would be beneficial to current techniques as it;

- Increases the number of counts by splitting the counts into count combinations, and
- Reduces the number of OD pairs associated with each count combination.

Both of the above qualities in the data improve the probability of deriving an optimal solution, matching the counts over distance and the trip matrix. With the OD pairs associated with each count over distance, the enhancement of the model's trip distribution will be a by-product of the process. With a network of ANPR cameras, the impact of potentially incorrect routing from the errors in the BPR curves could be nullified since the counts over distance are independent of route between the ANPR cameras but provided that a sample of the OD trips pass through the ANPR stations. This could be a significant improvement on current matrix estimation practices using traffic counts.

Whilst this work uses the ANPR data from the GFIP freeway network, which is limited in terms of the overall Gauteng road network, the methodology is expandable to include ANPR cameras deployed throughout the network. The fact that the 42 ANPR camera locations on the GFIP freeways provided approximately 340 logical G2G count combinations is noteworthy. The expansion and use of this process is certainly possible in other cities with wider coverage of the ANPR capable cameras throughout the network. For example, in the city of London there are 1 517 ANPR cameras throughout the city. (Transport for London, 2014). The combinations of potential G2G counts from this system over the overall network could be in the tens of thousands. This is a "big data" count-over-distance database that, if used in the estimation of trip matrices, could significantly improve current practices in the calibration of transport models.

7 CONCLUSIONS

The aim of this work was to evaluate the potential to improve traffic models using large-scale Automatic Number Plate Recognition (ANPR) data. The focus of many researchers and organisations has been to derive travel time and routing information from the traffic moving through the network using this technology. This is achieved from recording vehicles passing the entry/exit points of a closed cordon or along corridors. The derivation of origin destination information from this data has been limited for various reasons, ranging from availability of the suitable data and where it has been used, it has been limited to the ODs between measuring points and to freeway ramps in a closed environment.

This work explores the potential of using ANPR data collected from observation points (toll gantries) in an open layout, i.e. not necessarily around a cordon and can be at locations along one or more strategic routes throughout the network. The ANPR data derived from the Gauteng Freeway Improvement Project's open road tolling system constitutes such a layout. SANRAL made the ANPR data available for this project, together with the project's traffic model, used to derive the traffic forecasts for this toll road project.

The ANPR data provides data covering a significant spatial area connected by 201km of freeway within an overall network comprising approximately 2 000 km of alternative routes. The data comprises approximately 75 million entries per month relating to approximately 36 million trips. The data used in this study does not require expansion factors to make it relevant or determine hourly traffic volumes, daily flow profiles etc., which potentially introducing modelling errors at an early stage of the model development.

The processing of the data through the tracking of vehicles from one gantry to another provides various data sets usable in traffic model development, including journey times between gantries, traffic counts at each gantry and counts of vehicles that travel between specific gantry to gantry (G2G) pairs. From the comparison of the ANPR data to other data sources, the following is concluded:

- ANPR data provides accurate traffic counts. The ORT system classifies these counts by vehicle class for tolling purposes; this is not necessarily applicable to standard ANPR cameras. Comprehensive traffic observation (CTO) stations are permanent electronic counters with more locations throughout the network than the current ANPR coverage. They are also located on main roads on the alternative road network. These CTO counters measure vehicles and speeds for each lane at interchanges providing a

more diverse data set. For providing pure traffic counts, the standard electronic traffic counters may provide better information.

- Journey times from the ANPR are available for all times of the day and for each vehicle class. This data is accurate and accounts for all vehicles on the roads between ANPR camera locations (gantries). Unfortunately, the data is limited to the freeway network. In a network of ANPR camera, the derivation of accurate journey times for a multitude of routes is possible. Journey speeds/times are also effectively measured using probe data. A relatively small sample (5%-15%) in the traffic stream will accurately record journey times. Probe data vehicles are not restricted to any particular route, therefore providing information over the entire network and not just on specific routes. For the provision of travel time information, probe data may be the better choice.
- The number of ANPR count-over-distance (G2G) measurements increase significantly with the number of recording installations, whereas standard electronic counts yield one count per station. This disaggregation of the traffic counts into counts over distance is a significant advantage over standard electronic counts as “embedded” in the ANPR counts-over-distance is the traffic distribution.

Accepting the ANPR data as accurate and comprehensive, this data has the potential to enhance a traffic model’s trip matrices significantly especially if incorporated into trip matrix estimation processes. This potentially offers a significant improvement to the current methods of matrix estimation due to the ability to provide more independent and lower volume counts with each G2G count being associated with fewer OD pairs. According to (Ortúzar & Willumsen, 1998), this is recommended for improved matrix estimation.

The above indicates that ANPR data is not necessarily the answer to a modeller’s data needs, especially in terms of traffic counts and journey time measurements. The use of ANPR technologies to develop the trip matrix distribution has focused on closed systems, monitoring the movement along freeway networks. Expanding this work to an open layout and maximising its potential in improving a model’s trip distribution was a key element of this work. Noting that the G2G counts are not ODs and represent only a part of the trip and do not provide the actual origins or destinations, it is however concluded that by limiting the geographical areas from where the trips originate or are destined for, a significant improvement to the distribution of trips in the trip matrix could be achieved.

The methodology developed in this study enables the isolation of the trips within the cells of a model’s trip matrix that make up the G2G counts by processing select link matrices from the modelled links on which the ANPR recording devices are located. The process results in select-link-to-select-link (SL2SL) matrices and the total number of trips in these matrices equates to

the G2G count. The comparison of the measured G2G counts and modelled SL2SL outputs enables the validation (or otherwise) of the modelled distribution of trips through all ANPR locations.

Comparing the processed ANPR data from 2015 to the GFIP traffic model's 2015 forecasts has shown that the current GFIP model's 2015 forecasts, whilst in general appearing to have a reasonable correlation (within 10% of the light vehicle traffic counts), it requires re-calibration when viewed in terms of the webTAG criteria.

The comparison of the various elements of the data and the model highlighted the following issues:

- The recalibration of the freeway volume delay curves. However, it is preferable to use permanent traffic counters that measure speed and volume at specific locations for this purpose. The resultant journey times along the freeways should be validated against the ANPR data. The validation of travel times on the alternative road network will require probe data from other sources.
- The modelling of the heavy vehicles can be significantly improved in terms of:
 - The volumes of heavy vehicles on the network. The heavy vehicle volumes on the alternative routes must also be checked to ensure that it is not a diversion problem, in which case, the value of time may need to be re-assessed.
 - The freeway journey times for heavy vehicles can be checked and if necessary introduce differential speeds for trucks in the modelled network.
 - The distribution of heavy vehicles in the current model was based on somewhat experimental assumptions. With the improved ANPR distribution, the heavy vehicle trip matrices can be adjusted according to the G2G counts.

Whilst the comparison of modelled and actual traffic counts and journey times may meet general calibration criteria, the assessment of the G2G counts over distance has revealed that there are significant discrepancies in the composition of trips within the traffic stream in terms of the distribution. This was highlighted where the traffic volumes on critical links were modelled to a reasonable degree of accuracy but in one location there was half the actual number of short distance trips and at another it was twice the measured values.

Therefore, in terms of using large scale ANPR data to improve traffic models, the following conclusions are drawn:

- As a single data source, the ANPR data provided offers numerous elements of the data needed to develop transport models

- There are alternative data sources that can provide more comprehensive data for each traffic model's data requirement, for example counters that measure speeds or probe data that has much better network penetration
- The power of the ANPR data is in the ability to disaggregate it into counts over distance, which maximises its usefulness in terms of matrix estimation.
- Issues related to using the ANPR data within the POPI Act impacts on the availability of this data for general use and anonymising the data could make this more available.

8 RECOMMENDATIONS FOR FURTHER RESEARCH

This work has demonstrated the ability to extract traffic counts from ANPR locations that are specific to camera locations and routes through a series of camera. This data has the potential to provide a large number of counts over specific distances, a data set that has not been used to date.

The methodology to extract trips from cells in the overall model's trip matrices that make up these counts has been demonstrated in this work, and considered a "proof of concept". The process is expandable enabling modellers to improve on the standard matrix estimation techniques. This process would make a significant contribution to the transport modelling industry.

In addition, this process relies on the Select Link matrices taken from links where the ANPR camera are deployed. These select link matrices relate to the "prior" trip matrix and may not be complete or have numerous zero cells. The SATURN matrix estimation process has the option to seed zero cells but limit the extent to which cells are factored. The current matrix estimation process is an iterative one dealing with each traffic count in turn and repeating the process until convergence criteria are met. Research into the iterative nature of this process and appropriate convergence criteria will need further investigation.

The CPU requirements to process the ANPR data generated from all the select link matrices for all user classes, as in the GFIP model, is significant. On a standard computer, the run times can take 8 to 10 hours. Applying a greater spread of camera location and larger model matrices will result in significantly longer run times and require the use of a much more powerful CPU using multi-core processors. This consideration must be researched if this matrix estimation is developed further.

The issue with using the ANPR data within the POPI Act must be resolved. A centralised anonymising protocol should be developed where number plates are randomised in the same way during the same time periods irrespective of where the data is sourced. This will ensure the transferability of data between service providers and allow the analysis of long distance trips that may take place over several days.

9 REFERENCES

- Asakura, Y., Hato, E., & Kashiwadani, M. (2000). Origin-Destination Matrices Estimation Model Using Automatic Vehicle Identification Data and Its Application to the Han-Shin Expressway Network. *Transportation : Planning - Policy - Research - Practice*, 27(4), (pp. 419–438).
doi:10.1023/A:1005239823771
- Bain, R. (2009). *Credit Risk Analysis; Toll Road Traffic & Revenue Forecasts; An Interpreter's Guide*. Seville: Publicaciones Digitales.
- BlueScan. (2016). City Monitoring. Zurich, Switzerland. Retrieved January 16, 2016, from http://www.bluescan.org/english/downloads/citymonitoring_e.pdf
- Cambridge Systematics Inc. (2012). *Travel Time Data Collection*. Florida, USA: Florida Department of Transport, District IV.
- Carpenter, C., Fowler, M., & Adler, T. j. (2012). Generating Route Specific Origin-Destination Tables Using Bluetooth Technology. *Transportation Research Board No 2308*, 96-102.
doi:10.3141/2308-10
- Cascetta, E. (1984, Aug-Oct). Estimation of Trip Matrices from Traffic Counts and Survey Data: A Generalized Least Squares Estimator. *Transportation Research Part B Methodological*, pp. 289-299 18B.
- Castillo, E., Menéndez, J. M., & Sánchez-Cambronero, S. (2008). Traffic Estimation and Optimal Counting Location Without Path Enumeration Using Bayesian Networks. *Computer-Aided Civil and Infrastructure Engineering*, 23(3), , 189-207.
- COTO. (2013). *TMH 14 South African Standard Automatic Traffic Data Collection Format*. Pretoria: South African National Roads Agency Limited.
- COTO. (2013). *TMH 17 South African Trip Data Manual Version 1.01*. Roads Coordinating Body (RCB), Committee of Transport Officials. South Africa: South African National Roads Agency Limited.
- Department for Transport. (2014). *TAG Unit M3.1 Highway Assignment Modelling*. Department for Transport, Transport Appraisal and Strategic Modelling Division. London: Department for Transport.

- Department of Transport. (1995). *South African Trip Generation Rates - Second Edition, Report RR 92/228*. Pretoria: BKS Inc.
- Dixon, M. P., & Rilett, L. R. (2005). Population Origin-Destination Estimation Using Automatic Vehicle Identification and Volume Data. *Journal of Transport Engineering* 131(2), (pp. 75-82).
doi:10.1061/(ASCE)0733-947X(2005)131:2(75)
- Doblas, J., & Benitez, F. (2005). An Approach to Estimating and Updating Origin-Destination matrices Based oUpon Traffic Counts Preserving the Prior Structure of a Survey Matrix. *Transport Research Part B* 39, 565-591.
- Ermentrout, B. (2002). *Simplifying and reducing complex models*. Pittsburg, PA 15260: Department of Mathmatics, University of Pittsburg.
- Fisk, C. (1989). Trip Matrix Estimation From Link Traffic Counts: The Congested Network Case. *Transportation Research Part B Vol 23B No. 5*, 311-336.
- Gauteng Province. (2014). *Gauteng Province Household Travel Survey*. Roads and Transport. Retrieved from
[http://www.roadsandtransport.gpg.gov.za/services/Documents/2014%20Gauteng%20Household%20Travel%20Survey%20Report%20\(Provincial\).pdf](http://www.roadsandtransport.gpg.gov.za/services/Documents/2014%20Gauteng%20Household%20Travel%20Survey%20Report%20(Provincial).pdf)
- Google. (2009, August 25). The bright side of sitting in traffic: Crowdsourcing road congestion data. Retrieved June 2016, from <https://googleblog.blogspot.co.za/2009/08/bright-side-of-sitting-in-traffic.html>
- Hadavi, M., & Shafahi, Y. (2016). Vehicle Identification Sensor Models for Origin-Destination Estimation. *Transportation Research Part B*, 89, pp. 82-106.
- Hartgen, D., & San Jose, E. (2009). *Cost of Trip rates of Recent Household Travel Surveys*. The Hartgen Group, Charlotte NC. Retrieved June 2016, from
http://www.hartgengroup.net/Projects/National/USA/household_travel_summary/2009-11-11_Final_Report_Revised.pdf
- Hartgen, D., & San Jose, E. (2009). *Cost of Trip rates of Recent Household Travel Surveys*. The Hartgen Group, Charlotte NC. Retrieved June 2016, from
http://www.hartgengroup.net/Projects/National/USA/household_travel_summary/2009-11-11_Final_Report_Revised.pdf

- Hazelton, M., & Gordon, A. (2002). Estimation of Origin-Destination Trip Matrices from Link Counts. *European Transport Conference*. Cambridge: Association for European Transport .
- Highways England. (2016, September). *Research at Highways England*. Retrieved from Gov.UK: <https://www.gov.uk/government/organisations/highways-england/about/research#post-opening-project-evaluation-pope>
- Himayounfar, A., Ho, A., Zhu, N., Head, G., & Palmer, P. (2011). Multi-Vehicle Convoy Analysis Based on ANPR Data. *IET Conference*. Stevenage: The Institution of Engineering and Technology.
- Horn, R. A., & Johnson, C. R. (2012). *Matrix Analysis*. Cambridge University Press.
- Kirby, H. (1979). *Partial Matrix Techniques, Working Paper 111*. University of Leeds, Institute of Transport Studies. Retrieved from White Rose Research Online: <http://eprints.whiterose.ac.uk/2413/>
- Minguez, R., Sanchez-Cambronero, S., Castillo, E., & Jimenez, P. (2010). Optimal Traffic Plate Scanning Location for OD Trip Matrix and Route Estimation in Road Networks. *Transportation Research Part B*, 44(2), (pp. 282-298). doi:10.1016/j.trb.2009.07.008
- Ortúzar, J., & Willumsen, L. (1998). *Modelling Transport, Second Edition*. Wiley.
- Pollard, T., Taylor, N., & van Vuren, T. (2013). Comparing the Quality of OD Matrices: In Time and Between Data Sources. *European Transport Conference*. Frankfurt: Association for European Transport. Retrieved from <https://abstracts.aetransport.org/paper/index/id/10/confid/1>
- Pollard, T., Taylor, N., & van Vuren, T. (2013). Comparing the Quality of OD Matrices: In Time and Between Data Sources. *European Transport Conference*. Frankfurt: Association for European Transport. Retrieved from <https://abstracts.aetransport.org/paper/index/id/10/confid/1>
- Ramirez, S., Kovacic, K., & Ivanjko, E. (2013). *Origin-Destination Matrix Estimation of Traffic Flow on Highway Network*. University of Zegrab, Department of Intelligent Transport Systems, . Croatia: Faculty of Transport and Traffic Sciences.
- Robinson, A. (2014). The Use of Automatic Number Plate Recognition (ANPR) Data for Transport Model Development. *Australian Institute of Traffic Planning and Management (AITMP) National Conference*. Adelaide.
- Skrobanski, G., Logie, M., Black, I., Fearon, J., & Gilliam, C. (2010). OD Data Fusion. *European Transport Conference*. Glasgow: Association for European Transport Contributors.

Standard & Poor's. (2005). Traffic Forecasting Risk Study Update 2005: Through Ramp-Up and Beyond.

Statistics South Africa. (2016). *Statistical Release P0302: Mid-year Population Estimates*. Pretoria: Government Press.

Sun, Y., Zhu, H., Zhou, X., & Sun, L. (2014). VAPA: Vehicle activity patterns analysis based on Automatic Number Plate Recognition System Data. *2nd International Conference on Information Technology and Quantitative Management, ITQM*. Moscow, Russia: Elsevier B.V.

Tamin, O., & Willumsen, L. (1989). Transport Demand Model Estimation From Traffic Counts. *Transportation (1986-1998)*, vol. 16, no.1, (p. 3).

Transport for London. (2014, December 10). *WhatDoTheyKnow*. Retrieved from My Society: https://www.whatdotheyknow.com/request/tfl_anpr_cameras_and_data

van Vliet, D. (2015). *SATURN User Manual*. Leeds University, ITS. ATKINS.

van Vuren, T., & Carey, C. (2011). Building Practical Origin-Destination (OD/Trip) From Automatically Collected GPS Data. *European Transport Conferencr*. Glasgow: Association for European Transport Contributors.

van Vuren, T., Clarke, P., & Davidson, P. (2004). Building OD Matrices from a Wide Variety of Data Sources - Example of the West Midlands Strategic Transport Model. *European Transport Conference*. Strasbourg: Association for European Transport.

Van Zyl NJW, O. J. (July 2001). South African Experience with the Estimation of Values of Time from Stated Preference Studies and Their Use in Toll Road Models. *South African Transport Conference*. Pretoria.

Willumsen, L. (1981). Simplified Transport Models Based on Traffic Counts. *Transportation : Planning - Policy - Research - Practice*, 10(3), 257-278.

Wu, G. (2013, August 7). *Why More Data and Simple Algorithms Beat Complex Analytics Models*. Retrieved March 2016, from DataInformed: <http://data-informed.com/why-more-data-and-simple-algorithms-beat-complex-analytics-models/>

APPENDICES

TABLE OF APPENDICES

Appendix A: ANPR Data.....	1
Appendix B: Gantry to Gantry Distances	3
Appendix C: Data Miner Software	4
Appendix D: Preparation of Gantry Data	7
Appendix E: ANPR Gantries and Electronic (CTO) Count Stations.....	38
Appendix F: Comparison of ANPR and CTO Data.....	40
Appendix G: Modelling Process for Formula Validation	43
Appendix H: Gantry / Model Comparisons	49

Appendix A: ANPR Data

The ANPR data was obtained from the Gauteng Freeway Improvement Project (GFIP) Open Road Toll systems.

The extent of the GFIP freeway network is provided in Figure provided here.

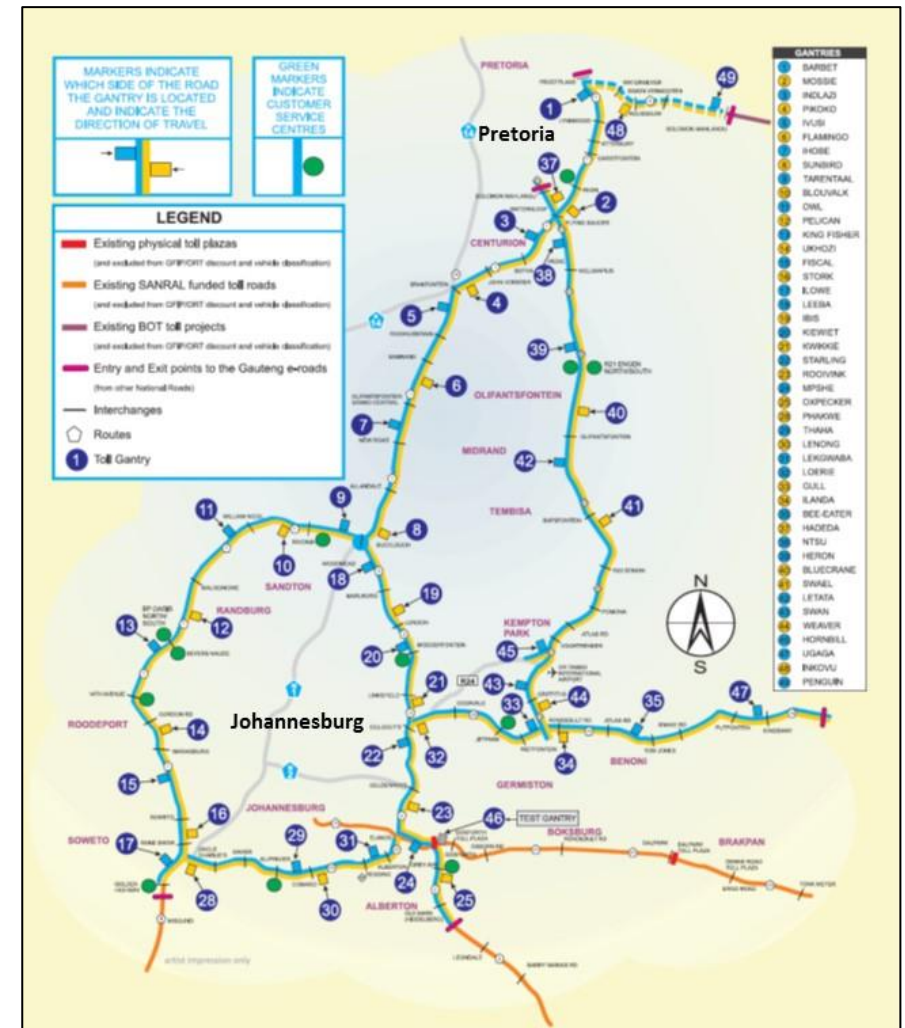


Figure 1: GFIP Project and Gantry Locations

Information in the extracted ANPR data records includes the following:

- The date and time the vehicle passed under the gantry
- The gantry number
- The vehicle classification
- The VLN ID

Extent of data used in this study – July 2015 data which comprises 77.8 million entries.

A sample of this data is provided here.

Date/Time	Gantry	Class	Vehicle ID
01/07/2015 09:43:26	1005	2	2081614
01/07/2015 09:43:30	1033	2	1410886
01/07/2015 07:29:38	1002	2	1400967
01/07/2015 07:29:01	1001	2	897839
01/07/2015 07:29:36	1018	2	4954274
01/07/2015 07:47:01	1005	2	2176313
01/07/2015 07:47:09	1002	2	2236412
01/07/2015 07:51:43	1020	2	2618336
01/07/2015 08:10:24	1023	2	3285196
01/07/2015 09:41:36	1019	2	5159164
01/07/2015 09:41:30	1011	2	4260223
01/07/2015 09:42:44	1007	2	4516063
01/07/2015 09:40:35	1008	2	468435
01/07/2015 09:41:44	1023	2	967466
01/07/2015 09:43:43	1021	2	5158981
01/07/2015 07:30:47	1040	2	2044439
01/07/2015 07:29:17	1019	2	1983836
01/07/2015 07:30:02	1011	2	248455
01/07/2015 07:47:32	1020	2	3638530
01/07/2015 07:48:58	1030	3	2314760
01/07/2015 07:48:01	1008	2	4954251
01/07/2015 08:10:24	1022	2	248372
01/07/2015 08:10:31	1020	2	2608085
01/07/2015 08:10:44	1002	2	3901860

Appendix C: Data Miner Software

The Data Miner software was developed in Java© and used for this study under a temporary licence.

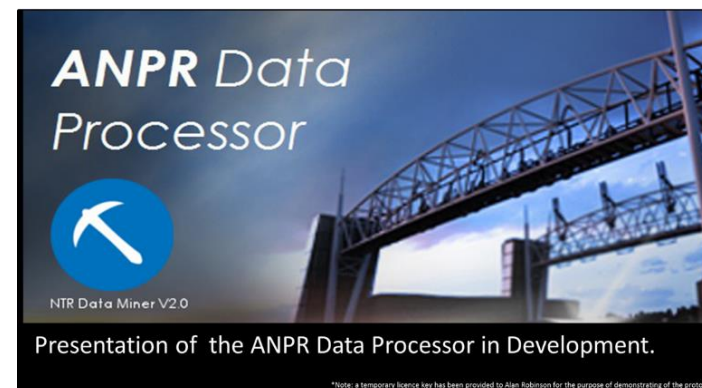
Stage 1: Data sorting and counting

The initial processing of the data is to sort the data by VLN ID, and then by time.

In the sorted data (right) the following can be deduced:

- VLN 2366 entered the freeway network between gantry 10 and gantry 12 (either at the William Nicol or Malibongwe Interchange), passed gantry 12, then 14 and then 16 before leaving the GFIP network (refer to the table to the right). The time taken travelling between gantry 12 and 16 is 10 minutes and 8 seconds.
- VLN 2373 entered the freeway network between gantry 11 and 9 (Bayers Naude or Malibongwe Interchange) and left the network either before Olifantsfontein (N1 northbound) or London Road (N3 southbound))
- VLN 2376 uses the N12 and N3 on two days.

During this process the vehicle counts by gantry and vehicle class are recorded in 15-minute time bins.



Date/Time	Gantry	Class	VLN ID
02/07/2015 06:09:11	1012	2	2366
02/07/2015 06:15:37	1014	2	2366
02/07/2015 06:19:19	1016	2	2366
02/07/2015 06:59:38	1009	2	2373
01/07/2015 06:11:48	1031	2	2376
01/07/2015 06:19:51	1022	2	2376
01/07/2015 06:26:36	1020	2	2376
02/07/2015 06:15:17	1031	2	2376
02/07/2015 06:24:00	1022	2	2376
02/07/2015 06:32:21	1020	2	2376
01/07/2015 05:43:09	1002	4	2408
01/07/2015 05:55:14	1040	4	2408
01/07/2015 08:26:02	1040	2	2453
01/07/2015 08:32:11	1041	2	2453

Stage 2: Preparation of trip files

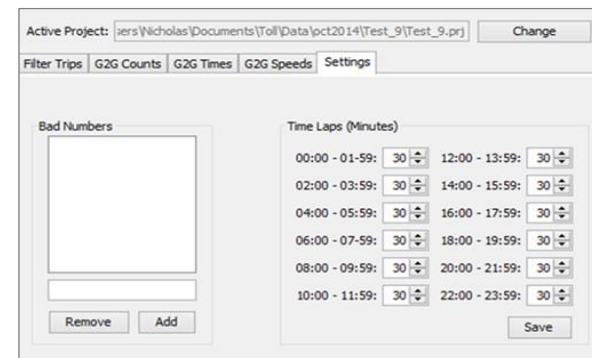
The processing of the sorted data is dependent on two elements:

- Cleaning the data, and
- The definition of trip records based on the output requirements in terms of vehicle class, days of the week and time period

The VLN IDs relating to missing and cloned number plates are entered as “Bad Numbers” and excluded from further analysis.

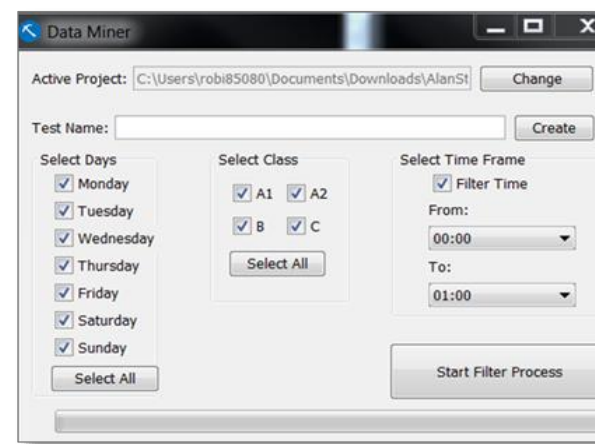
VLNID	Class	Start Passage Time	End Passage Time	Day Of Trip	Start Tolling Point	End Tolling Point	Travel Time Sec	Tolling Point Count
13235	2	01/07/2015 06:37	01/07/2015 06:37	Wednesday	1031	1031	0	1
13235	2	01/07/2015 11:49	01/07/2015 12:01	Wednesday	1024	1018	691	4
13235	2	01/07/2015 14:05	01/07/2015 14:19	Wednesday	1009	1023	839	4
13235	2	01/07/2015 16:47	01/07/2015 16:47	Wednesday	1024	1024	0	1
13238	2	01/07/2015 09:28	01/07/2015 09:28	Wednesday	1010	1010	0	1
13238	2	01/07/2015 15:00	01/07/2015 15:11	Wednesday	1012	1014	708	2
13245	2	01/07/2015 07:22	01/07/2015 07:47	Wednesday	1002	1006	1485	3
13245	2	01/07/2015 16:59	01/07/2015 17:23	Wednesday	1005	1001	1415	3
13246	2	01/07/2015 05:49	01/07/2015 05:53	Wednesday	1024	1022	246	2
13246	2	01/07/2015 08:15	01/07/2015 08:18	Wednesday	1023	1025	192	2
13246	2	01/07/2015 08:53	01/07/2015 08:53	Wednesday	1024	1024	0	1
13246	2	01/07/2015 17:32	01/07/2015 17:48	Wednesday	1021	1025	957	3
13255	2	01/07/2015 09:44	01/07/2015 09:54	Wednesday	1007	1003	588	3
13255	2	01/07/2015 10:20	01/07/2015 10:20	Wednesday	1001	1001	0	1
13255	2	01/07/2015 11:30	01/07/2015 11:30	Wednesday	1006	1006	0	1
13255	2	01/07/2015 12:41	01/07/2015 12:41	Wednesday	1008	1008	0	1
13260	2	01/07/2015 05:42	01/07/2015 05:54	Wednesday	1034	1020	706	3
13260	2	01/07/2015 06:18	01/07/2015 06:31	Wednesday	1019	1025	745	4
13260	2	01/07/2015 14:18	01/07/2015 14:26	Wednesday	1024	1020	498	3
13260	2	01/07/2015 15:06	01/07/2015 15:27	Wednesday	1019	1047	1279	5
13269	2	01/07/2015 14:36	01/07/2015 14:36	Wednesday	1007	1007	0	1
13276	2	01/07/2015 15:32	01/07/2015 15:32	Wednesday	1023	1023	0	1
13281	2	01/07/2015 19:53	01/07/2015 19:57	Wednesday	1023	1025	240	2

Maximum times in minutes are entered for each two-hour time period to define the end of a trip between successive gantries.



Trip files are produced for each test according to the selected attributes, i.e. days of the week, vehicle class and time period.

Stage 3:



Production of G2G information

Using the start and end gantry (tolling point) and the start and end time the trips associated with each logical gantry to gantry pair are counted and entered into a square matrix. The sums of entries are divided by the number of entries to derive the averages, i.e. there are say 4 Mondays or 23 week days or 31 days if only Mondays, all weekdays or all days were selected respectively.

The G2G matrices of vehicles and times are produced for each hour of the times selected.

Additional outputs include gantry traffic counts and journey times between gantries.

Appendix D: Preparation of Gantry Data

Traffic Counts: Average weekday February 2015

Gantry	1				2				3				4				5				6			
	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total
0	171	30	53	254	206	25	47	277	189	23	35	247	172	21	33	225	228	29	42	298	223	25	47	295
1	99	30	55	183	119	22	42	183	101	24	40	165	96	20	27	143	132	32	46	210	126	33	56	214
2	91	35	48	174	98	23	38	159	79	23	33	136	75	22	28	125	98	29	39	166	99	28	41	169
3	131	35	53	220	137	26	38	202	109	26	36	171	104	22	28	154	127	31	43	201	135	27	38	200
4	318	47	86	451	639	42	53	734	321	35	48	403	393	36	34	463	301	47	60	409	507	37	51	596
5	1098	70	133	1300	4153	64	97	4314	1148	54	68	1271	3283	41	46	3370	1167	85	95	1346	4617	88	78	4782
6	3341	118	147	3606	8623	138	118	8878	4629	100	77	4806	5159	79	57	5295	4577	156	93	4826	8135	119	64	8318
7	4193	160	127	4480	8540	138	94	8773	7166	133	66	7365	4484	106	53	4643	6255	174	75	6504	7904	136	79	8120
8	3650	251	164	4065	6813	250	137	7200	5201	211	93	5505	4742	172	76	4990	4900	251	105	5256	7153	210	94	7457
9	3353	278	187	3818	5017	237	138	5392	4197	259	107	4563	4036	180	82	4298	4599	335	129	5063	5693	266	113	6072
10	3188	270	170	3629	4467	255	150	4871	4006	264	95	4366	3542	186	85	3813	4515	326	114	4955	4818	262	124	5204
11	3217	261	164	3642	4468	266	157	4891	4051	247	98	4396	3520	195	87	3802	4505	299	113	4917	4817	274	126	5216
12	3451	245	164	3859	4654	290	161	5105	4314	238	95	4647	3618	212	90	3920	4648	278	112	5038	4852	294	133	5278
13	3748	235	164	4147	4689	299	175	5162	4660	232	96	4988	3585	212	96	3893	5005	276	114	5395	4600	289	135	5024
14	4087	237	160	4484	4859	330	186	5375	4984	230	92	5306	3662	231	94	3987	5320	269	109	5697	4406	303	135	4845
15	5009	235	157	5401	5608	343	191	6143	6407	210	88	6704	4500	236	95	4831	6788	236	102	7125	4875	310	128	5313
16	6922	227	151	7299	7159	318	190	7667	7452	171	77	7699	5891	218	100	6210	7941	203	88	8232	5192	277	120	5590
17	6176	199	171	6547	6019	264	192	6474	6829	158	91	7078	5235	167	98	5499	7224	190	90	7504	4791	209	113	5113
18	3983	133	169	4285	3577	163	166	3905	4878	90	80	5048	3206	109	89	3405	5232	120	86	5439	3240	127	107	3474
19	2311	84	140	2535	2214	105	150	2469	2622	54	64	2741	1885	64	75	2024	2867	78	72	3018	2143	87	86	2315
20	1536	66	127	1729	1524	77	129	1731	1611	44	60	1715	1199	47	65	1311	1845	60	71	1975	1479	64	81	1623
21	1164	56	116	1336	1227	59	109	1394	1194	38	57	1290	927	38	54	1019	1329	47	67	1443	1127	55	74	1256
22	845	48	89	982	847	45	87	979	875	34	50	958	662	31	46	740	965	41	57	1063	787	41	57	885
23	491	40	70	601	487	34	66	587	524	29	40	593	385	26	37	448	578	34	46	659	489	30	45	563
Total	62572	3390	3065	69027	86143	3811	2909	92863	77547	2928	1685	82160	64360	2670	1575	68606	81145	3627	1968	86741	82206	3591	2125	87922

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

Gantry	7				8				9				10				11				12			
	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total
0	275	33	43	351	319	28	48	394	303	24	25	351	281	22	24	327	169	23	22	214	232	18	22	273
1	159	29	39	227	178	36	54	268	169	22	23	214	156	25	25	206	94	19	21	134	128	18	25	170
2	122	29	39	189	133	32	46	211	119	22	23	164	113	25	25	163	76	21	21	118	89	20	24	132
3	142	39	49	230	179	32	40	251	151	24	26	201	137	24	22	183	109	22	23	155	96	20	22	138
4	332	55	68	455	624	51	56	731	548	28	29	605	298	32	31	361	482	30	29	541	183	23	25	231
5	1347	90	121	1558	4515	141	94	4749	2841	59	41	2942	1568	68	47	1683	3468	71	42	3582	811	44	34	889
6	4353	170	107	4630	11448	221	89	11758	6925	79	35	7039	5048	140	54	5241	7092	69	36	7197	2522	92	38	2651
7	5531	190	86	5807	10923	257	95	11275	6959	87	37	7084	5216	173	53	5442	5991	75	44	6110	3317	113	41	3471
8	4809	288	146	5243	8726	373	117	9216	6331	165	62	6558	4261	271	65	4597	5388	172	59	5618	3247	191	61	3499
9	4972	416	175	5564	7635	453	148	8236	6062	266	72	6400	4496	338	80	4913	4850	254	61	5165	3161	241	70	3472
10	4941	403	154	5498	6702	428	158	7289	5382	277	72	5731	4631	314	70	5015	3999	246	63	4308	3014	223	63	3299
11	5013	360	147	5521	6562	434	157	7154	5113	293	74	5480	4882	307	69	5258	3803	263	67	4133	3127	216	65	3409
12	5285	341	145	5771	6691	462	160	7313	5146	302	74	5521	5136	289	70	5496	3752	259	70	4080	3305	217	66	3588
13	5570	343	144	6057	6700	469	171	7340	5185	296	76	5558	5304	296	72	5671	3721	247	66	4034	3556	221	67	3843
14	5968	333	139	6440	6548	480	162	7190	5156	314	74	5545	5573	288	61	5922	3629	273	71	3973	3788	229	61	4077
15	7909	310	129	8349	7709	462	155	8327	6229	321	77	6627	6631	233	58	6922	3973	282	72	4327	5276	216	62	5553
16	8460	249	103	8812	8494	386	143	9023	6422	301	68	6791	6746	159	49	6955	4326	249	64	4639	6716	184	62	6963
17	7822	243	114	8178	7378	307	137	7822	5452	242	66	5760	6530	131	53	6714	3780	181	60	4022	6392	165	67	6624
18	6278	174	112	6564	5100	186	120	5406	4266	141	56	4463	5018	80	47	5145	2804	106	50	2960	4725	90	53	4868
19	3536	102	90	3728	3309	120	94	3522	2868	82	45	2995	3467	55	39	3561	1988	59	41	2088	2751	47	40	2837
20	2326	78	87	2491	2256	86	84	2426	1905	55	40	2001	2312	45	36	2393	1369	43	36	1448	1608	36	35	1678
21	1696	60	76	1832	1662	69	79	1810	1477	42	37	1555	1729	37	32	1798	1071	36	32	1140	1188	27	32	1247
22	1205	52	69	1325	1168	53	61	1282	1182	36	34	1252	1213	33	27	1273	777	31	29	836	873	24	26	924
23	700	40	57	796	673	39	47	758	791	27	31	849	706	29	24	759	456	25	26	507	552	20	23	596
Total	88751	4425	2441	95616	115631	5605	2513	123750	86980	3507	1196	91683	81451	3415	1132	85998	67166	3057	1106	71329	60655	2693	1084	64433

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

Gantry	13				14				15				16				17				18			
	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total
0	136	20	23	179	179	17	25	220	160	18	25	203	169	19	24	212	143	24	40	207	179	28	34	240
1	83	16	22	121	99	16	25	139	90	17	24	131	96	16	25	137	97	22	39	159	94	27	33	153
2	64	17	22	103	77	17	24	117	75	17	22	114	83	17	24	123	84	20	36	140	77	29	33	139
3	91	20	23	134	96	18	25	139	87	20	25	132	107	19	24	151	129	28	44	200	98	32	37	167
4	346	32	30	409	262	25	27	315	253	35	37	325	278	27	27	331	441	51	66	557	283	39	60	382
5	2679	64	55	2797	1177	41	47	1265	1828	69	66	1963	1108	43	48	1199	3035	112	110	3257	1604	78	104	1785
6	7587	118	69	7774	3360	73	39	3473	7372	134	87	7593	3228	76	40	3344	6328	104	93	6525	5756	158	95	6009
7	5884	125	66	6075	4408	108	52	4567	7161	145	81	7387	4074	103	49	4226	4711	100	87	4899	6249	225	93	6568
8	4461	201	69	4731	3844	190	75	4109	4447	220	94	4761	3291	170	66	3527	3061	119	98	3278	5323	387	140	5850
9	3281	213	63	3556	2961	225	84	3271	2890	224	90	3204	2680	202	73	2955	2345	148	105	2597	4490	460	149	5099
10	2801	203	64	3067	2523	218	80	2820	2476	218	89	2784	2330	205	79	2614	2051	172	117	2340	4094	405	130	4630
11	2731	203	68	3002	2406	220	84	2709	2448	213	89	2751	2232	206	85	2523	1935	175	122	2232	4150	357	127	4634
12	2750	205	68	3023	2466	219	80	2765	2619	211	89	2919	2324	212	87	2623	1955	180	122	2257	4142	330	122	4595
13	2732	204	65	3000	2615	227	87	2929	2730	208	85	3023	2486	211	90	2786	1956	184	117	2257	4152	328	116	4596
14	2820	230	70	3120	2848	236	83	3167	2977	216	87	3280	2725	208	87	3021	2018	204	118	2340	4300	316	107	4723
15	3173	247	77	3496	3759	245	83	4086	3584	223	86	3893	3591	228	87	3906	2173	225	116	2514	4892	281	94	5267
16	3862	206	67	4135	5790	233	84	6108	4734	185	73	4992	5590	212	87	5889	2170	217	117	2504	5471	214	82	5768
17	3407	144	57	3607	6103	231	93	6428	4177	133	59	4369	5969	204	97	6270	2057	178	114	2349	4860	149	73	5081
18	2209	83	52	2344	4087	130	72	4290	2703	83	55	2841	3943	106	75	4124	1815	134	100	2049	3489	95	64	3648
19	1431	51	44	1525	2279	61	51	2391	1682	57	51	1789	2222	57	56	2335	1420	93	102	1615	2234	69	64	2367
20	1033	33	37	1104	1280	39	42	1361	1198	38	41	1277	1349	37	45	1431	1101	64	87	1252	1577	56	61	1694
21	772	26	32	830	880	28	37	945	935	29	35	999	949	31	39	1019	847	49	72	968	1169	47	51	1267
22	550	22	31	604	616	24	32	672	619	23	32	674	652	25	35	711	565	41	62	667	810	37	47	894
23	304	20	28	352	401	19	26	446	372	21	28	422	402	22	27	451	311	36	48	395	454	33	40	527
Total	55185	2703	1202	59090	54515	2859	1358	58731	57618	2757	1450	61825	51876	2657	1374	55907	42749	2680	2130	47558	69948	4181	1955	76084

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

Gantry	19				20				21				22				23				24			
Time	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total
0	226	29	37	291	195	26	33	254	215	30	46	290	215	40	84	338	232	36	76	344	138	25	48	211
1	122	29	38	189	96	27	32	155	128	29	43	200	136	33	75	243	157	38	67	262	86	22	45	153
2	93	28	36	156	82	25	31	138	104	30	43	178	123	37	75	235	131	38	64	234	80	26	49	155
3	126	29	32	187	108	29	36	173	148	31	40	219	179	56	89	323	161	43	64	268	112	32	66	210
4	555	35	39	629	338	37	62	437	586	43	46	675	664	75	169	908	311	50	98	459	348	42	123	514
5	2053	73	60	2186	2341	70	108	2519	1974	77	68	2119	3235	152	304	3691	1139	90	157	1386	1903	81	216	2201
6	5134	113	62	5308	7920	153	101	8174	5223	140	80	5442	7870	223	248	8342	3512	179	172	3863	5860	157	215	6231
7	6689	168	57	6914	7560	191	101	7852	7555	218	74	7846	6772	261	276	7309	4467	286	168	4921	4802	167	204	5173
8	5531	257	87	5875	6423	386	155	6964	5706	313	105	6125	5247	452	364	6062	3979	395	215	4589	2735	199	220	3154
9	5269	353	116	5738	5318	468	157	5943	5199	389	141	5728	4861	579	373	5813	3852	423	262	4537	2156	215	215	2586
10	5101	365	134	5600	4836	410	136	5382	5065	407	164	5636	4838	566	356	5760	3800	414	286	4499	1983	200	200	2383
11	5055	396	139	5589	4898	385	136	5419	5018	423	173	5613	4842	558	365	5766	3982	410	308	4699	1990	197	180	2367
12	5158	397	132	5687	4876	369	132	5376	5075	412	159	5647	4857	541	356	5754	4170	404	317	4890	2021	187	170	2377
13	5402	410	146	5959	4731	367	127	5225	5281	422	171	5873	4729	526	331	5586	4487	409	325	5220	2116	193	157	2465
14	5464	415	140	6019	4748	370	111	5228	5371	436	167	5974	4807	520	317	5645	4926	411	328	5665	2173	192	141	2507
15	6827	406	140	7373	5083	329	100	5511	6528	414	169	7110	4896	488	292	5675	6195	396	334	6926	2310	195	140	2646
16	7013	335	118	7465	5450	250	85	5786	6882	317	135	7334	5344	420	266	6030	8271	314	292	8877	2761	162	124	3047
17	6394	319	134	6847	5057	187	79	5322	6626	326	155	7107	4937	366	264	5566	8175	280	303	8758	2550	126	112	2788
18	4903	199	117	5219	3574	111	68	3753	5127	224	136	5486	3205	233	231	3669	6289	237	344	6870	1792	94	106	1992
19	2822	109	80	3012	2349	73	68	2490	2949	109	89	3147	2192	142	210	2544	3413	134	256	3803	1224	65	95	1385
20	1756	71	65	1892	1693	56	62	1811	1835	66	73	1974	1691	101	190	1981	2159	99	210	2468	895	49	88	1032
21	1316	59	62	1437	1277	44	52	1372	1351	59	75	1484	1248	79	156	1483	1520	78	182	1780	678	43	75	795
22	964	51	50	1066	896	36	47	979	955	55	62	1071	869	67	131	1067	1092	63	155	1309	463	34	62	559
23	568	36	41	645	494	29	40	564	543	38	52	633	511	59	105	675	594	48	108	750	297	31	55	383
Total	84541	4682	2062	91285	80343	4427	2057	86827	85442	5005	2465	92912	78265	6573	5627	90465	77013	5273	5089	87375	41476	2733	3106	47315

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

Gantry	25				28				29				30				31				32			
Time	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total
0	124	25	43	192	245	28	50	323	111	19	31	161	113	20	35	168	126	24	42	192	309	45	82	437
1	88	23	40	151	159	30	45	234	67	19	31	116	73	21	32	126	78	22	41	141	184	52	68	304
2	82	24	37	142	127	32	46	205	59	18	32	109	61	23	29	113	70	20	40	130	161	46	64	270
3	115	27	36	178	140	37	44	221	78	19	32	129	73	22	32	127	114	24	41	179	178	47	68	293
4	204	31	52	287	280	38	64	383	242	30	40	311	162	29	49	240	434	36	74	544	461	62	101	625
5	534	47	79	660	1077	68	98	1244	1195	50	64	1309	661	45	74	781	2229	68	144	2441	2459	114	173	2746
6	1383	80	86	1549	3825	139	112	4075	4079	80	65	4224	2610	98	95	2803	5820	88	126	6034	7244	226	181	7651
7	1617	103	74	1793	4787	219	116	5123	4259	85	65	4409	3951	202	111	4264	4988	113	133	5234	7334	292	173	7799
8	1235	127	95	1457	3760	304	135	4199	2893	144	85	3122	2779	250	118	3146	3946	192	161	4299	6852	511	221	7583
9	1141	140	127	1407	3062	308	135	3505	2110	154	84	2348	2218	224	112	2554	3174	223	157	3555	5698	570	251	6519
10	1114	148	138	1400	2784	277	131	3192	1877	164	93	2134	2066	198	106	2370	2853	235	168	3256	5481	543	272	6296
11	1165	149	151	1466	2792	257	132	3182	1830	164	99	2093	2108	184	105	2397	2723	242	177	3142	5628	528	288	6445
12	1205	151	166	1521	2991	251	134	3376	1901	174	103	2179	2237	179	103	2518	2726	246	183	3155	5663	511	294	6468
13	1345	152	166	1663	3169	240	136	3544	2012	177	104	2292	2357	176	102	2635	2795	249	178	3221	5527	524	287	6338
14	1562	167	183	1912	3616	250	134	4000	2130	185	104	2419	2619	184	100	2902	2829	258	181	3268	5571	526	282	6380
15	1939	183	194	2316	4617	248	127	4991	2532	209	109	2851	3178	174	97	3448	3066	273	174	3513	5671	478	272	6421
16	2988	166	185	3340	6227	208	118	6553	3442	213	110	3765	4613	145	92	4850	3679	272	162	4113	6056	341	223	6620
17	3677	163	197	4036	6686	187	113	6985	2990	172	111	3272	4555	118	85	4758	3267	217	159	3643	5729	280	238	6247
18	2777	135	216	3127	4831	148	133	5113	1902	104	90	2096	2966	83	94	3142	2225	134	147	2505	4375	209	239	4823
19	1697	84	166	1947	3176	107	113	3395	1282	64	76	1422	1548	59	80	1687	1528	83	127	1738	3017	148	210	3375
20	1086	60	136	1282	2033	82	91	2205	899	45	68	1012	1002	46	62	1110	1059	56	107	1221	2215	114	183	2512
21	725	56	110	891	1357	50	76	1483	671	32	56	760	706	30	56	792	801	41	88	931	1664	85	159	1909
22	508	47	96	651	914	43	72	1029	464	30	47	541	499	25	49	572	536	38	72	646	1226	71	141	1439
23	283	34	65	381	542	34	66	642	284	27	38	349	285	23	45	353	310	37	60	406	696	63	109	867
Total	28593	2319	2836	33747	63196	3584	2418	69198	39306	2379	1738	43423	43439	2559	1859	47857	51376	3190	2941	57506	89400	6384	4580	100364

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

Gantry	33				34				35				37				38				39			
	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total
0	115	25	39	179	105	22	42	169	121	24	39	184	46	14	19	79	111	21	38	171	112	22	38	172
1	69	22	37	128	71	21	35	126	69	26	37	131	26	14	19	58	60	24	38	122	60	24	37	121
2	60	25	37	122	66	19	35	119	67	30	35	132	24	14	19	56	54	28	32	115	57	28	31	115
3	83	32	40	155	100	22	40	162	84	34	41	159	36	13	19	68	70	25	39	133	71	24	41	136
4	204	40	58	302	327	31	54	412	203	39	59	301	178	15	22	215	194	32	69	294	180	33	74	286
5	753	70	84	907	1853	60	83	1996	653	67	82	802	765	26	25	816	848	52	99	999	713	54	99	866
6	1916	112	76	2103	6671	108	96	6875	1556	113	77	1745	1742	40	30	1812	3294	126	115	3536	2809	129	112	3049
7	2772	169	75	3016	6658	129	92	6879	2047	140	66	2254	2087	48	29	2164	4930	232	107	5269	3996	238	104	4337
8	2425	238	113	2776	3804	179	113	4096	1828	186	92	2105	1445	62	38	1545	3357	285	128	3769	2941	277	127	3345
9	2516	266	118	2899	2832	219	121	3173	1903	220	110	2233	1153	70	38	1261	2438	243	140	2821	2174	231	132	2536
10	2682	269	126	3077	2644	231	126	3001	1997	210	114	2322	1097	71	41	1208	2173	205	131	2508	1945	188	124	2256
11	2768	268	129	3166	2621	249	128	2998	2044	206	115	2365	1126	83	39	1248	2100	191	128	2420	1915	167	119	2201
12	2908	244	125	3276	2665	241	130	3036	2164	197	116	2478	1223	86	40	1350	2212	175	132	2519	1989	154	125	2269
13	3034	237	124	3395	2700	243	128	3071	2281	195	114	2589	1302	92	39	1433	2301	165	127	2593	2033	146	118	2296
14	3167	241	126	3534	2694	266	130	3090	2463	189	115	2766	1359	109	42	1510	2471	165	124	2760	2239	150	117	2506
15	3797	217	121	4135	2657	263	137	3057	2947	172	109	3228	1642	126	44	1812	2916	159	126	3201	2804	136	119	3060
16	5669	195	113	5977	2812	234	117	3163	4672	143	101	4916	2158	129	46	2333	3862	136	113	4112	3881	120	106	4107
17	5282	189	124	5595	2406	185	113	2704	4851	131	113	5095	1817	125	45	1987	3681	106	113	3899	3721	97	112	3929
18	3222	130	103	3455	1685	125	108	1917	2805	100	100	3005	1023	77	35	1135	2250	79	117	2445	2313	74	114	2501
19	1641	75	87	1803	1109	76	96	1281	1413	68	86	1567	614	40	32	686	1302	55	104	1461	1309	52	101	1462
20	1020	46	77	1143	774	58	89	921	920	44	75	1038	432	26	28	487	954	40	87	1081	957	39	86	1082
21	727	37	61	825	610	45	82	737	627	33	62	722	296	19	24	340	690	35	75	799	661	34	71	766
22	513	33	56	602	391	38	66	494	489	29	55	573	191	15	23	228	495	28	57	579	445	28	54	527
23	298	29	49	376	228	34	54	316	272	28	47	348	104	14	21	139	278	28	47	353	264	28	46	338
Total	47639	3208	2099	52946	48482	3096	2214	53792	38475	2622	1959	43057	21884	1328	755	23967	43039	2635	2285	47959	39589	2473	2203	44264

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

Gantry	40				41				42				43				44				45				47			
Time	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total	Light	Sm Hgv	Lg Hgv	Total
0	82	19	36	137	99	24	43	165	120	27	44	191	165	22	25	212	126	16	22	165	71	24	32	127	91	23	34	147
1	50	16	32	99	60	20	40	119	70	26	41	136	100	20	23	143	74	19	22	115	46	23	31	100	57	25	32	114
2	39	16	33	88	54	21	38	113	61	30	38	128	85	19	23	127	58	20	20	99	36	23	30	89	53	28	30	112
3	64	16	30	110	92	22	40	153	78	25	47	150	89	18	27	135	54	18	22	94	44	24	35	103	83	32	36	150
4	453	21	38	512	508	28	48	584	206	36	89	330	240	24	41	306	111	17	26	154	101	28	59	188	192	35	56	282
5	1428	39	68	1535	1698	61	81	1840	939	61	127	1127	701	38	46	785	443	27	33	503	403	53	91	547	654	60	80	794
6	4067	63	88	4218	5231	122	112	5466	3350	138	144	3632	2275	93	61	2429	1922	65	45	2033	1371	109	89	1568	1419	106	79	1604
7	4256	71	70	4396	5590	138	103	5831	4396	266	148	4810	3175	180	67	3422	2929	94	50	3073	2103	158	90	2350	1571	132	64	1767
8	2891	116	100	3106	3318	174	127	3619	3347	372	185	3903	2314	271	95	2680	2027	132	54	2213	1665	185	119	1969	1220	144	82	1445
9	2073	124	105	2302	2429	194	146	2769	2489	332	184	3005	1871	220	89	2181	1704	156	66	1925	1509	207	115	1831	1138	157	99	1394
10	1860	133	113	2106	2170	221	157	2548	2227	271	176	2673	1844	208	80	2133	1728	152	69	1949	1514	202	118	1834	1132	142	97	1371
11	1872	146	118	2136	2160	238	161	2559	2192	247	166	2605	1917	196	80	2193	1791	156	67	2015	1529	202	126	1857	1129	135	105	1369
12	1958	159	124	2241	2248	245	176	2668	2266	229	174	2669	2008	192	76	2276	1929	155	73	2158	1614	191	127	1932	1190	129	107	1425
13	2077	171	126	2374	2341	246	174	2762	2315	222	159	2697	2137	196	76	2408	1952	156	70	2177	1723	185	114	2023	1239	120	101	1460
14	2217	195	138	2550	2452	274	182	2908	2519	228	159	2906	2265	206	79	2551	2050	159	73	2281	1829	187	112	2128	1401	123	101	1626
15	2701	242	150	3093	2865	325	196	3385	3175	212	158	3544	2594	192	79	2865	2496	171	71	2737	2279	177	114	2570	1665	118	95	1878
16	3500	258	147	3906	3623	333	191	4148	4438	190	138	4766	3736	172	75	3983	4235	170	73	4478	3412	158	105	3674	2492	99	92	2683
17	3206	275	154	3636	3277	322	202	3802	4450	175	146	4771	3342	140	73	3555	3407	141	74	3622	3151	144	114	3409	3118	102	102	3321
18	1810	165	127	2102	1926	202	162	2290	2605	123	146	2874	1652	85	61	1797	1717	78	61	1856	1912	116	106	2134	1960	78	91	2130
19	1066	91	118	1274	1125	115	137	1377	1423	76	120	1620	1034	65	54	1153	998	45	49	1092	894	72	84	1050	1058	55	83	1196
20	742	59	100	901	789	80	113	982	994	52	100	1145	778	47	45	871	741	35	41	817	552	46	73	671	697	39	71	807
21	555	42	82	680	629	59	92	779	683	42	87	812	669	44	41	754	556	28	35	619	424	44	60	528	433	31	57	522
22	360	34	67	461	399	45	80	524	464	34	65	562	536	36	34	606	481	23	33	536	300	42	49	391	323	29	47	399
23	204	24	52	279	228	32	62	321	265	31	56	352	332	30	30	392	270	21	27	318	171	33	41	245	184	28	40	252
Total	39533	2496	2216	44244	45309	3542	2861	51713	45071	3443	2894	51408	35861	2713	1381	39955	33798	2056	1174	37028	28654	2631	2033	33318	24498	1967	1781	28246

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

All Vehicle Counts FROM: 6.00 to 7.00

To From	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47									
1	1360																																																							
2		3401		2772		303		316		52		20		6		7				17		4		63		18			6		7					24				269	1054			224			4									
3	854		1689																																	12				59	385			80			9									
4				775		81		131		22		8		3		1				10		7		4		2																														
5	217		405		1266																																				3	38														
6						1913		2853		428		190		60		58					239		244		128		57			3		10			8	4											5									
7	219		351		1052		753																																																	
8								3517		872		260		72		56				457		514		161		57			7		17			42		15								12		20										
9	54		105		185		126		1632											440		590		133		43				2			76		24									39		40										
10										677		212		45		47																																								
11	28		59		101		106		819		1235										336		323		9		5							48		11									38		17									
12												503		268		286											1																													
13	19		36		75		88		617		1260		1236								184		78		1																					13		5								
14														564		786																															4									
15	10		19		65		39		138		447		999		1009						5		1																																	
16																434																															6		2							
17	13		19		49		39		162		424		1087		1206		2295	2	5	18			210																																	
18	16		17		55		58			261		83		6		2		293																												1		29		8						
19																					263		268		75		30			3		6				45		17									20		35							
20	15		17		103		99			508		173		8		1		751																																						
21																																																								
22	38		13		87		86			156		11									433		417		1353																															
23														8		37																																								
24	36		12		103		122			218	4	5	108	1	221				618		541		964		2488																															
25																																																								
26																																																								
27																																																								
28					1				7		132		486		532																																									
29	4		1		6		4														26		27		227																															
30										1		30		258		324																																								
31	35		12		65		57														301		213		683																															
32			1		10		43							4		48					482		492																																	
33	24																																																							
34	32				2		25			255		47	2		36						462		440																																	
35																																																								
36																																																								
37	125				152		16		16		2										1																																			
38	71				167		16		18		3		1																																											
39	92				6																																																			
40																																																								
41																																																								
42	314				71																																																			
43	82				11																																																			
44																																																								
45	73				4																																																			
46																																																								
47																																																								

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

All Vehicle Counts FROM: 10.00 to 11.00

To From	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47					
1	1259																																																			
2		1611		836		412		653		146		68		19		39				55		41		42		24			1		4					10			166	560			86			6						
3	704		1357																																	4				33	129			21			2					
4				435		189		233		54		22		4		6				25		28		18		5					2																					
5	225		315		790																																				2	7										
6						820		1140		201		98		34		54				131		167		106		45			1		8			10	5									1		3						
7	333		477		1191		1043																																													
8							1806		455		204		55		49				288		347		154		51			2		15			46	24									13		14							
9	65		127		205		165		1240										395		468		102		37					6		72	34									44		21								
10										1104		419		55		48																3																				
11	44		63		104		112		638		849									268		273		28		15							54	28									32		16							
12												828		293		334								1		9					73		37																			
13	27		32		68		64		272		495		587							105		62		1		1							14	6									11		4							
14														475		670							37																					4								
15	7		4		19		13		26		91		201		364					1																																
16																500							32			25				78	88	6	4											4	5							
17	21		17		41		25		40		89		158		294		1150			2	3		114			28																		16		17						
18	23		26		61		82		222		84		11		3		305																																			
19																				620		616			150		40			5		14			99	45									53		30					
20	25		32		102		105			401		146		10		2						481		584																												
21																							1137		327		77			39		61			115	59			1	1			1		56		35					
22	43		23		88		88			137		17								248		246		1620										190	96				20	8			13		211		78					
23												10		48											917			425			79		130																			
24	27		12		54		52			71	2	5	36		102				131		112		411		935					127		140			18	4				7	3			4		47		6				
25																																																				
26																																																				
27																																																				
28					1		1		4		63		194		372																																					
29	3		1		4		4												9		8		118			55																										
30									1		23		111		213																																					
31	29		17		54		43													101		71		488			34																									
32	2		2		19		49						4		56					319		413			644		177			143		119		2445																		
33	12																																																			
34	18				3		11			94		33			12					102		114			85		8			43		26																				
35																																																				
36																																																				
37	67			54		25		38		9		5				2										4																										
38	78			57		26		33		10		3		1		1																																				
39	55			4																																																
40																																																				
41																																																				
42	351			58		2																																														
43	53			8																																																
44												2			1					7		8																														
45	67			3																																																
46																																																				
47																																																				

Appendix E: ANPR Gantries and Electronic (CTO) Count Stations

ORT gantries are located on freeway sections between interchanges. CTO stations are installed at interchanges and include the on-ramp(s) and the main through lanes between the off- and on-ramps. The sum of the ramps and through lanes provide the comparative traffic volumes to the downstream ORT gantry.

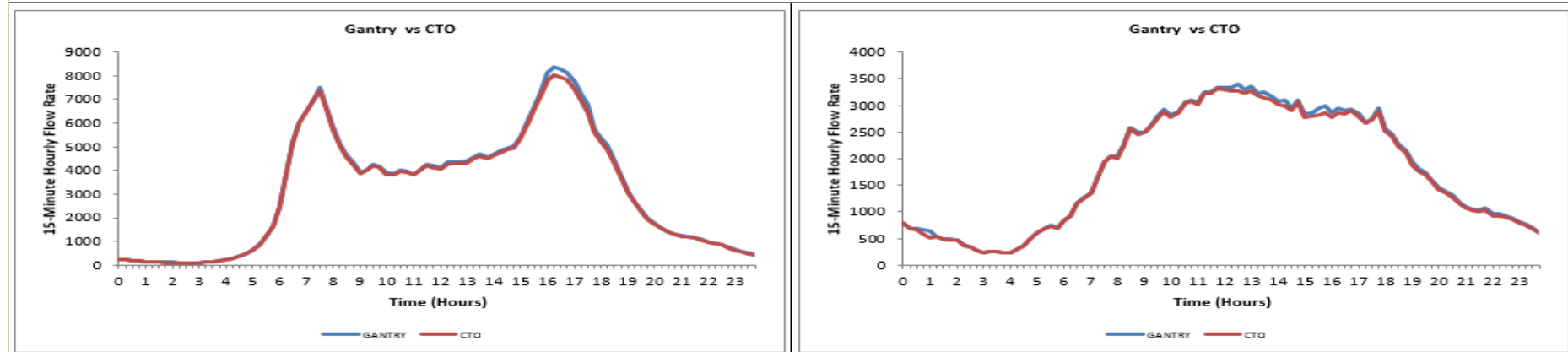
The gantry numbers and the corresponding CTO station number and location are provided in the Table to the right. These numbers are used in the following sample of data comparisons. In the graphs that follow the gantry numbers have a 1000 added to them, which is how they are referred to in the database.

The following graphs provide a sample of comparisons of the average weekday and weekend flow profiles. Gantry numbers 3, 5, 8 and 12 show very similar volumes with the gantry figures being marginally higher occasionally. Gantry numbers 18 and 22 show significant differences. These are due to the CTO data containing zeros where data is missing reducing the monthly average.

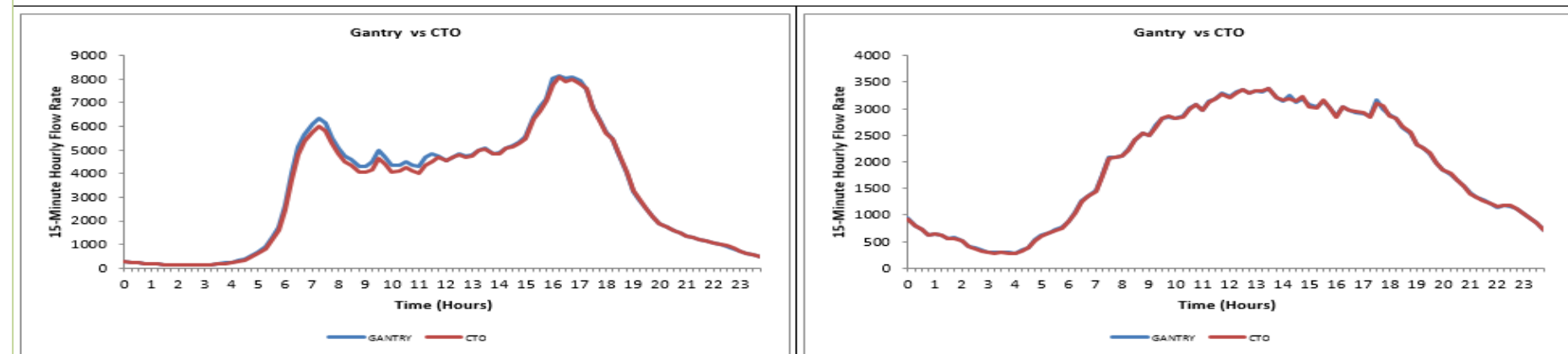
Gantry Number	CTO Station Number	SITE DESCRIPTION
1	1834	At Lynnwood Rd I/C NB On Ramp
2	1843	At Rigel Ave SB On Ramp
3	1846	At Botha Ave NB On Ramp
4	1849	At John Vorster Dr SB On Ramp
5	1851	At N1 & N14 Northbound Diverge (NB Only)
6	1858	At Olifantsfont Rd Southbound Off Ramp (SB Only)
7	1859	At New Rd Northbound On Ramp (NB Only)
8	1863	At Allandale Rd Southbound On Ramp (SB Only)
9	1869	At Rivonia Rd EB On Ramp
10	1871	At Rivonia Rd WB On Ramp (WB Only)
11	1874	At Malibongwe Dr NB On Ramp (NB Only)
12	1875	At Molibongwe Dr SB On
13	1878	At 14th Avenue NB On Ramp
14	1881	At Gordon Rd SB On Ramp
15	1884	At N17 Soweto NB On Ramp
16	1885	At Soweto Highway SB On Ramp
17	1890	Golden Highway NB On NO DATA
18	1893	At Marlboro Dr NB On Ramp
19	1894	At Marlboro Dr SB On Ramp
20	1898	At Modderfontein Rd NB On Ramp (NB Only)
21	1901	At Linksfield SB On Ramp
22	1905	At Van Buuren Rd Northbound On Ramp (NB Only)
23	1938	At Rand Airport Rd SB Off Ramp (SB Only)
24	1940	At N3/N17 NB On Ramp
25	1921	At Grey Avenue SB On Ramp
28	1934	At N1N & N1S Diverge
29	1928	At Kliprivier Rd EB On Ramp (EB Only)
30	1945	At Reading I/C WB On Ramps (WB Only)
31	1942	At Reading I/C EB On Ramps (EB Only)
32	1957	At R24 & N12 Merge (WB Only)
33	1964	At M44 Jet Park Rd EB On Ramp (EB Only)
37	1907	At Hans Stirjdom Dr SB On Ramp
38	1910	At Nellmapius Dr NB On Ramp
39	1912	At R562 Olifantsfontein Rd NB On Ramp (NB Only)
40	1911	At Nellmapius Dr SB On Ramp
41	1915	At R25 Kaalfontein/Bapsfontein SB On Ramp (SB Only)
42	1914	At R25 Kaalfontein/Bapsfontein NB On Ramp (NB Only)
43	1952	At Griffith Rd NB On Ramp (NB Only)
44	1953	At Griffith Rd SB On Ramp (SB Only)
45	1950	At R24E & R21N Merge (NB Only)

Appendix F: Comparison of ANPR and CTO Data

GANTRY - CTO NUMBER	1003-1846
SITE DESCRIPTION (CTO)	At Botha Ave NB On Ramp
Weekday Mon Tue Wed Thu Fri	
Weekend Sat Sun	



GANTRY - CTO NUMBER	1005-1851
SITE DESCRIPTION (CTO)	At N1 & N14 Northbound Diverge (NB Only)
Weekday Mon Tue Wed Thu Fri	
Weekend Sat Sun	

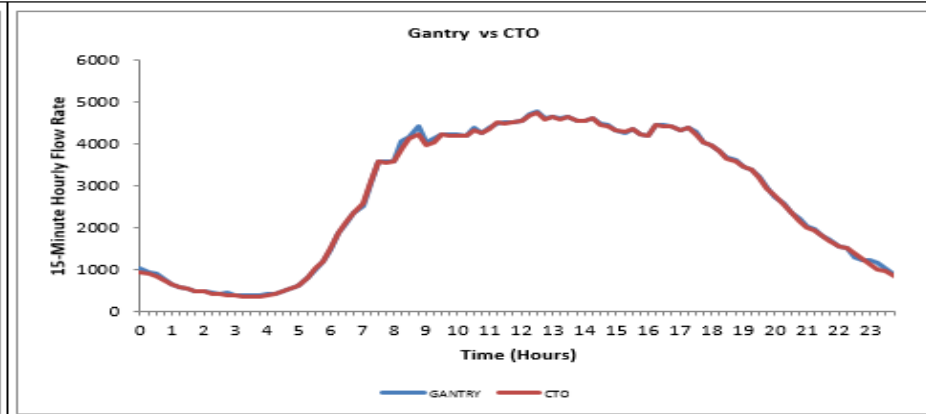
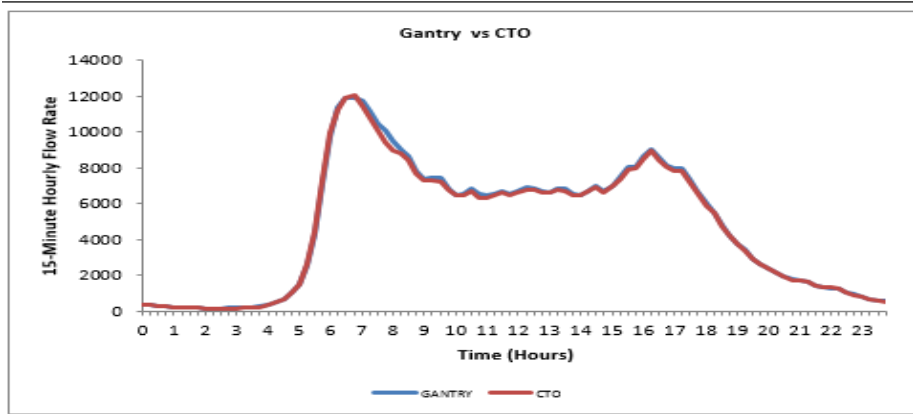


IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

GANTRY - CTO NUMBER 1008-1863
 SITE DESCRIPTION (CTO) At Allandale Rd Southbound On Ramp (SB Only)

Weekday
 Mon Tue Wed Thu Fri

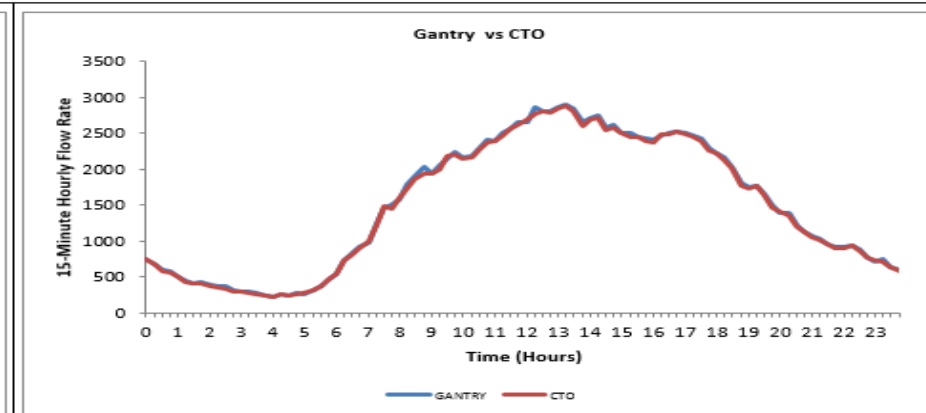
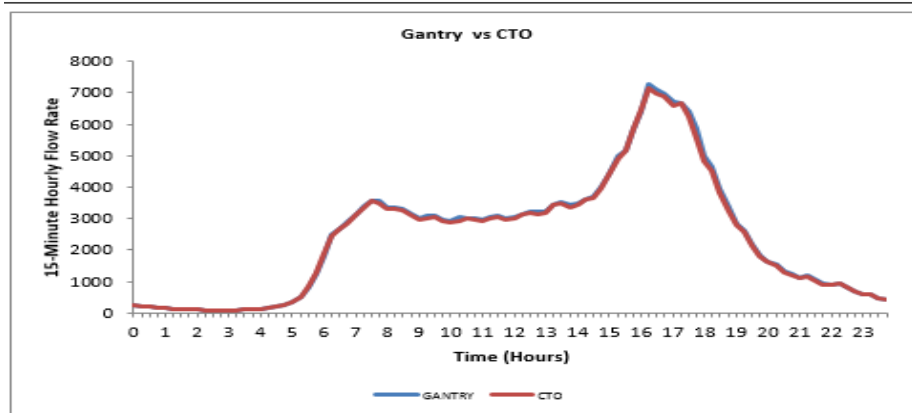
Weekend
 Sat Sun



GANTRY - CTO NUMBER 1012-1875
 SITE DESCRIPTION (CTO) At Mollibongwe Dr SB On

Weekday
 Mon Tue Wed Thu Fri

Weekend
 Sat Sun

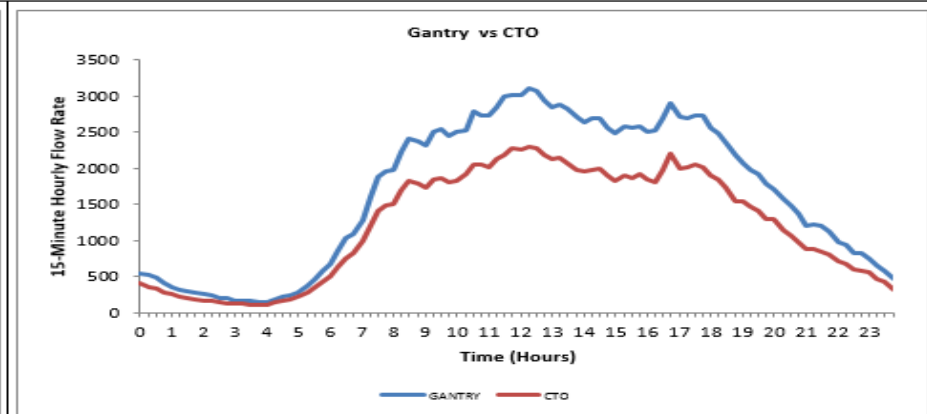
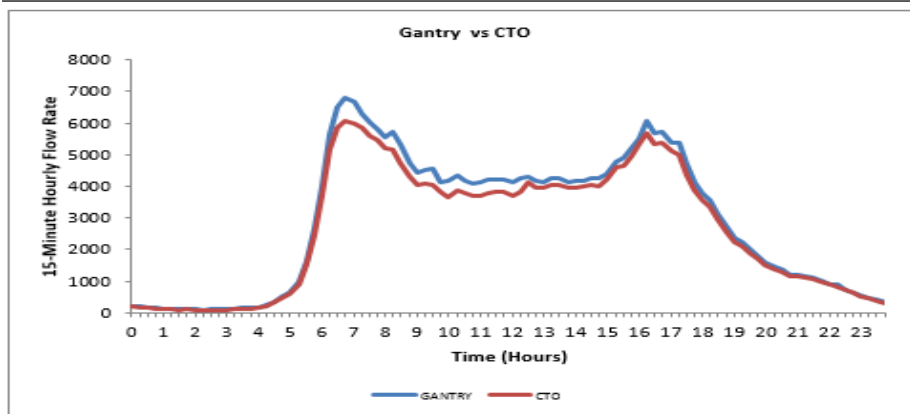


IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

GANTRY - CTO NUMBER 1018-1893
 SITE DESCRIPTION (CTO) At Marlboro Dr NB On Ramp

Weekday
 Mon Tue Wed Thu Fri

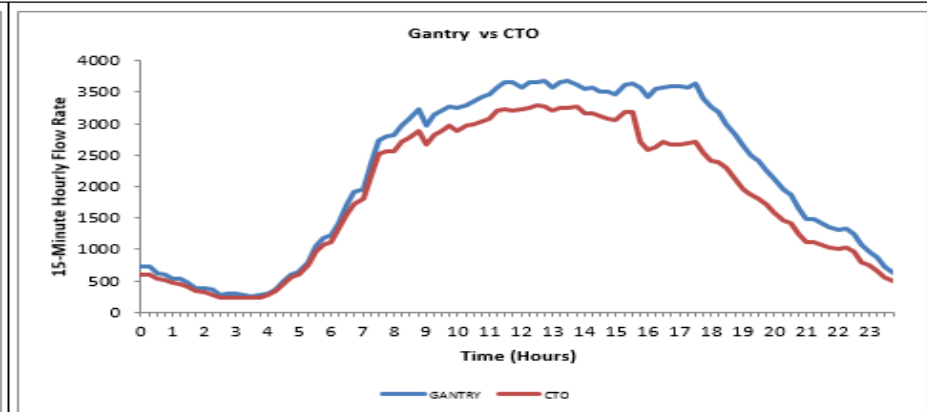
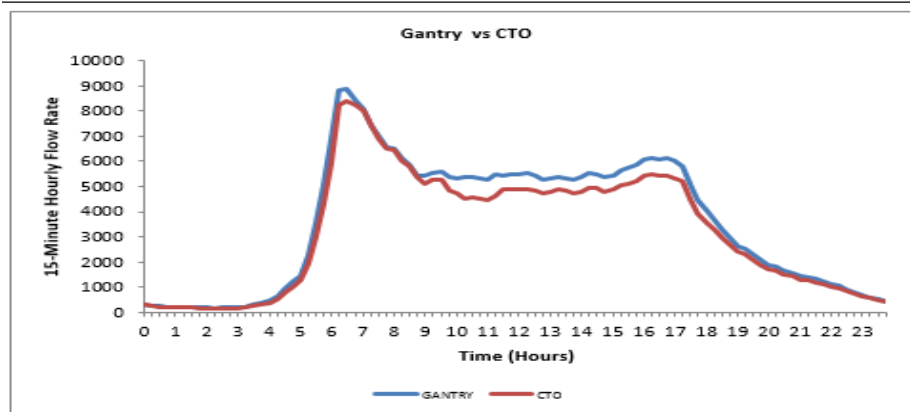
Weekend
 Sat Sun



GANTRY - CTO NUMBER 1022-1905
 SITE DESCRIPTION (CTO) At Van Buuren Rd Northbound On Ramp (NB Only)

Weekday
 Mon Tue Wed Thu Fri

Weekend
 Sat Sun



Appendix G: Modelling Process for Formula Validation

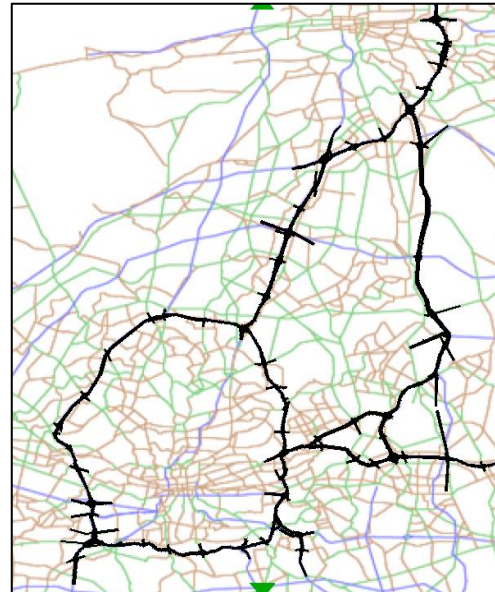
The Model

The freeway portion of the Gauteng (GFIP) 2015 morning peak hour forecast model was extracted using the SATURN cordon procedures. This results in the following files:

The freeway network including interchanges (see →)

The 2015 morning (AM) peak hour comprising six stacked user class matrices including:

- Light vehicle commuter trips
- Light vehicle business trips
- Light vehicle other trips
- Heavy vehicles (small) Class B
- Heavy vehicles (large) Class C

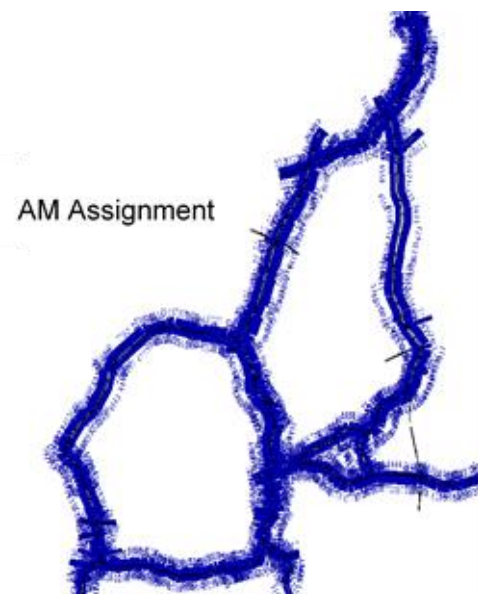


The testing of the G2G formula will be done using the light vehicles, i.e. the first three of the above matrices were added together and tested against the Class A2 data from the toll gantries. Without the option of route choice on the freeway network, averaging the VOT and VOC parameters will not alter the assignment outputs. Only the light vehicle matrices are assigned to the network.

Assignment

The light vehicle trip matrix (toll gantry Class A2) was assigned to the freeway network. The assigned volumes are depicted here →.

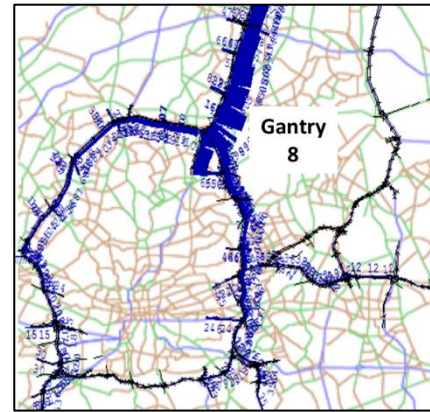
The assignment indicates the light vehicle flows on each section of the freeway network. The assignment also includes the trips entering and exiting the network through the interchanges as well as the traffic that passes over the freeways through the interchanges.



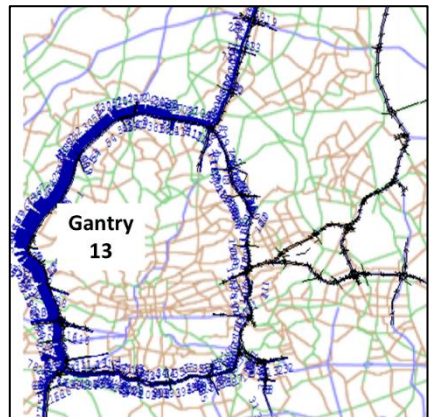
Select Link Matrices

Select link (SL) matrices are extracted from each of the 42 gantry locations.

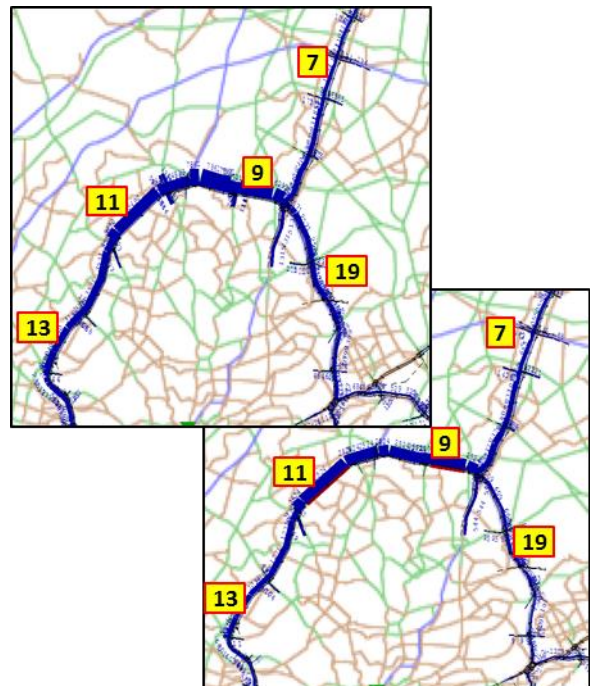
Two examples of the assigned SL matrices are provided here. The selected gantries are numbers 8 and 13. As shown, some trips pass a number of locations in addition to the selected ones.



As can be seen from these diagrams the traffic that passes through each gantry could have travelled over long distances on the freeway network. This means that some trips in the select link matrices from a single gantry location will pass through other gantry locations.



Likewise, the select link assignment below passed through two gantry locations; 11 and 9. The upper assignment is for trips that pass through at least one of the gantries and the lower assignment are for trips that pass through both locations. In the first assignment trips enter and exit the freeways through the interchanges between the gantry locations. In the second assignment this does not happen but trips do continue beyond the next up- and downstream gantry locations.



Gantry Associations

The G2G matrix formula requires the select link matrices of the two gantry locations under question and the upstream and downstream gantries. Where there is more than one entry point, i.e. through system interchanges, there are instances where there can be up to three upstream or downstream gantries.

An Association Table of all G2G combinations whereby the upstream and downstream gantries are recorded was created. For each gantry these upstream and downstream gantries were obtained from the map (Figure). For example, Gantry 4 is located on the southbound carriageway of the N1 to the south of the R21 Interchange. Upstream gantries include number 2 on the N1 north of the R21 Interchange, number 37 on the R21 southbound north of the N1 and number 38 on the R21 northbound south of the N1. Downstream of gantry 4 is gantry 6 on the N1 southbound between the Samrand and Olifantsfontein Interchanges.

This can be expanded to identify specific trip movements for a network monitored by surveillance cameras with ANPR capability and a model that represents the monitored area.

Identify required G2G Matrices

Only certain gantry combinations are considered probable trips. Whilst not impossible, other combinations of gantries are highly unlikely and are probably “lost souls” or the result of misread or possibly more than one vehicle with similar number plates. The non-zero cells and the diagonals of the G2G Distance matrix provide the combinations of the G2G matrices that can be produced.

A list of all combinations of gantries was created and the upstream and downstream associated gantries were paired with the “origin” and “destination” gantry respectively. The

Gantry	Upstream			Downstream		
	1	2	3	1	2	3
1	3	38	37	0	0	0
2	0	0	0	4	40	0
3	5	0	0	1	40	0
4	2	37	38	6	0	0
5	7	0	0	3	0	0
6	4	0	0	8	0	0
7	9	18	0	5	0	0
8	6	0	0	10	19	0
9	11	0	0	7	19	0
10	8	18	0	12	0	0
11	13	0	0	9	0	0
12	10	0	0	14	0	0
13	15	0	0	11	0	0
14	12	0	0	16	0	0
15	17	28	0	13	0	0
16	14	0	0	29	0	0
17	0	0	0	15	29	0
18	20	0	0	7	10	0
19	8	10	0	21	0	0
20	22	32	0	18	0	0
21	19	0	0	23	33	45
22	24	31	0	20	33	45
23	21	32	44	25	30	0
24	0	0	0	22	30	0
25	23	31	0	0	0	0
28	30	0	0	15	0	0
29	16	17	0	31	0	0
30	23	24	0	28	0	0
31	29	0	0	22	25	0
32	0	0	0	20	23	43
33	21	22	0	43	0	0
37	0	0	0	1	4	40
38	39	0	0	1	4	0
39	42	0	0	38	0	0
40	2	3	37	41	0	0
41	40	0	0	20	23	44
42	43	45	0	39	0	0
43	32	33	0	42	20	22
44	41	0	0	20	23	0
45	21	22	0	42	0	0
46	0	0	0	0	0	0
47	0	0	0	0	0	0

gantry numbers were substituted with the select link matrix file names relative to the respective gantry location. From this list a Windows batch file was created calling the matrix manipulation programs (MX) with the relevant list of select link matrices. Each time the program is called, the subroutine (KEY File) calls eight matrices, but the program can only call in a matrix once. Therefore, if the eight matrices contain either the same origin and destination matrix or there are only two up and down stream matrices with the remaining matrix containing zero, a means to create eight separately named matrices was included in the Batch File.

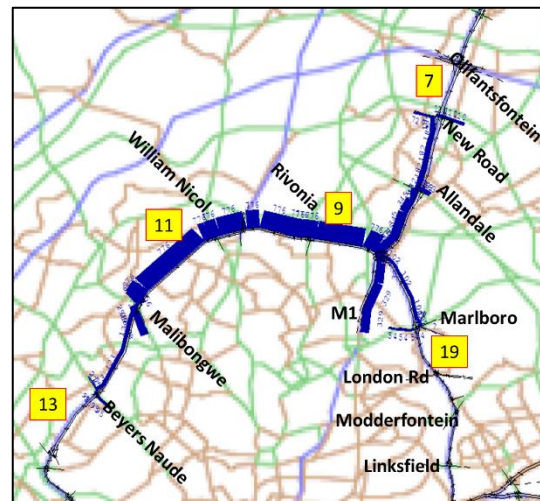
A total of 482 G2G matrices are created, which represent all realistic gantry combinations. These were exported to csv format for input into Excel and into a corresponding G2G matrix for comparison with the actual data.

To demonstrate the output matrices from the G2G formula whereby it is expected that trips only pass through the selected gantries (and those in between where applicable), the following examples are used:

- Gantry 8 only
- Gantry 11 to 9
- Gantry 13 to 7
- Gantry 8 and 23
- Gantry 2 and 25

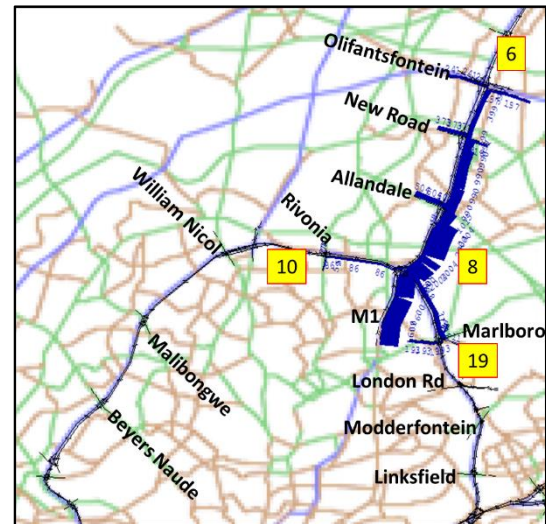
Gantry 8 Only

The assignment of the matrix of trips that only pass through gantry 8 shows that all trips enter the freeway after gantry 6, i.e. at the Olifantsfontein, New Road or Allandale Road interchanges and none of the trips continue past gantries 10 and 19, i.e. at either Rivonia Road or Marlboro Interchanges and a large portion exit the GFIP network on the M1 south



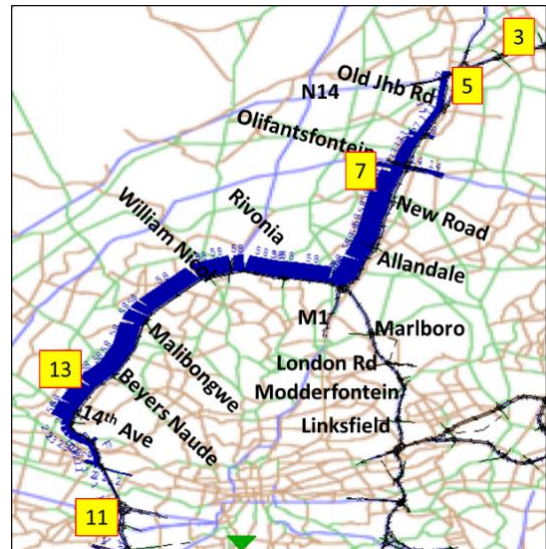
Gantry 11 and 9

Trips that pass through gantry 11 and 9 only enter the GFIP network at either Beyers Naude I/C or Malibongwe Interchanges, i.e. before gantry 13 and exit the freeways before the London Road Interchange (before gantry 19) and the Olifantsfontein Interchange (before gantry 7)



Gantry 13 and 7

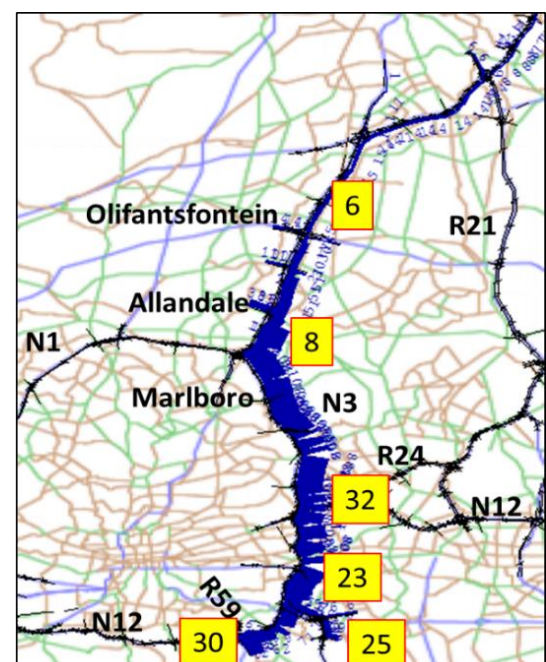
Trips that pass through gantry 13 and 7 enter the freeway between gantry 15 and 13 (Maraisburg to 14th Ave Interchanges) and exit between gantry 7 and 5 (Olifantsfontein and Old Johannesburg Rd Interchanges)



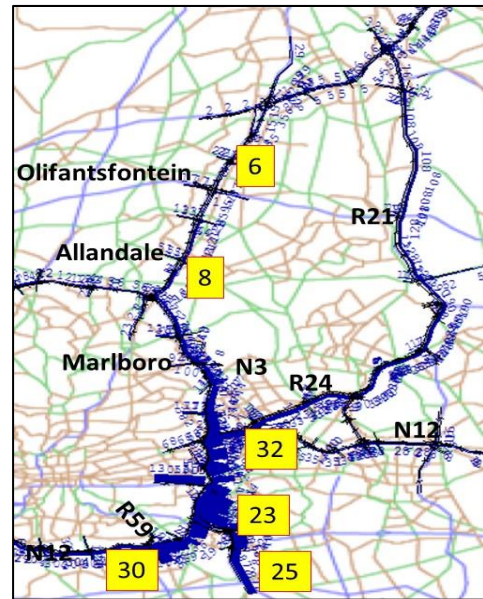
Gantry 8 and 23

The trips from gantry 8 on the N1 north of the Buccleuch Interchange to gantry 23 north of the Elands Interchange should only enter from the Olifantsfontein Interchange and leave before the Reading (R59) Interchange on the N12 and the Old Barn Interchange on the N3.

In this instance, the assignment shows trips from the N1 north of the Olifantsfontein Interchange.



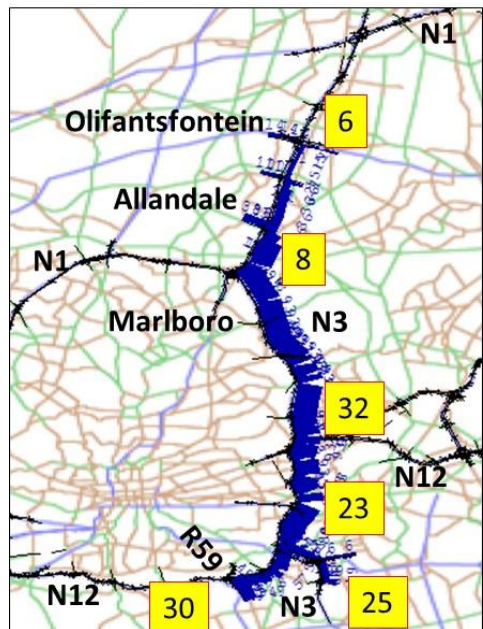
A select link analysis, from the freeway cordon model, through the N3 at the Gantry 23 location show that there are multi-route trips from the north along the N1/N3 and the R21/R24/N3. This multi-routing results in “leakage” past Gantry 8. These trips enter the system through Gantry 34 on the N12 to the east of the Gillooly’s Interchange.



This results shows that there is a need to ensure that any potential leakage through alternative routes must be plugged by identifying the gantry location(s) between the selected start and end gantries (in this case 8 and 23). Adding gantry 32 to the list of up- and downstream gantries in the formula “plugs” the leak and only includes the trips between gantry 8 and 23.

Adding the additional potential leakage gantries to the formula can be achieved by adding gantries 32, 33 and 45 to any path that involves trips that travel north/south on

the N3. This is a relatively easy task in this model but may be substantially more intricate on the full model.

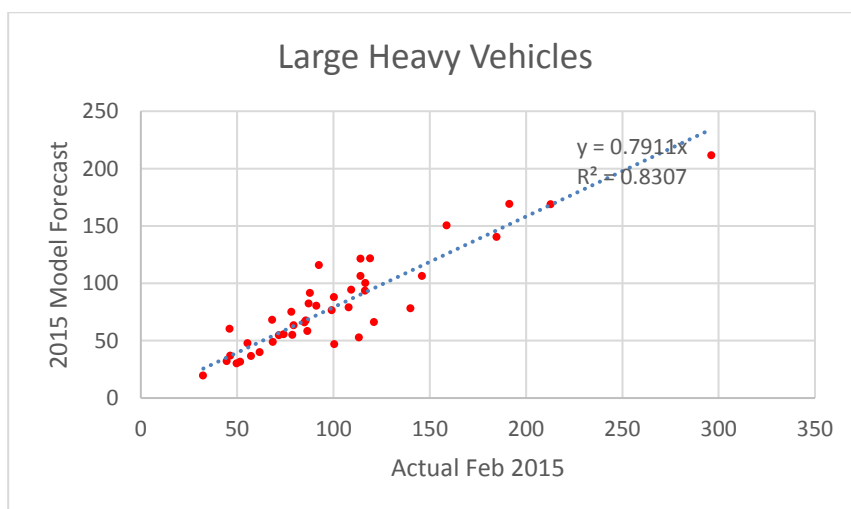
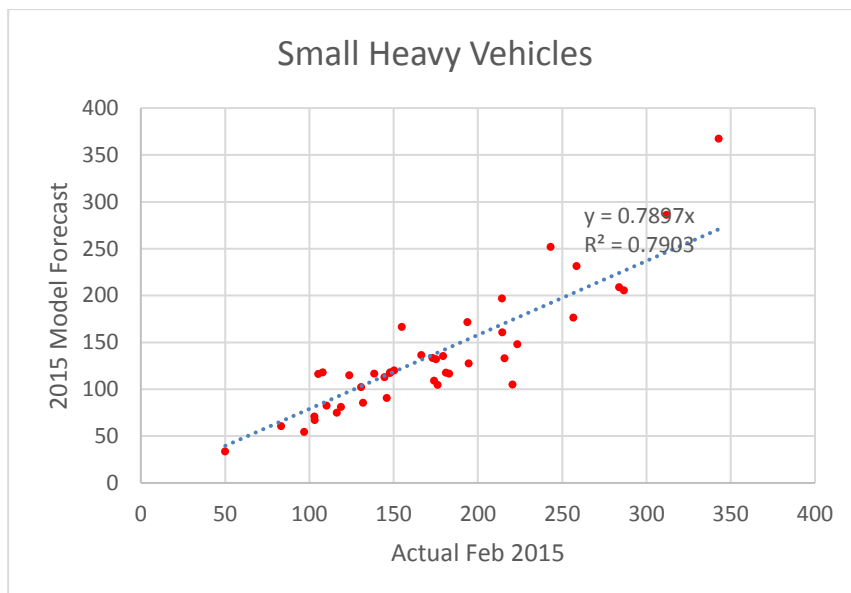
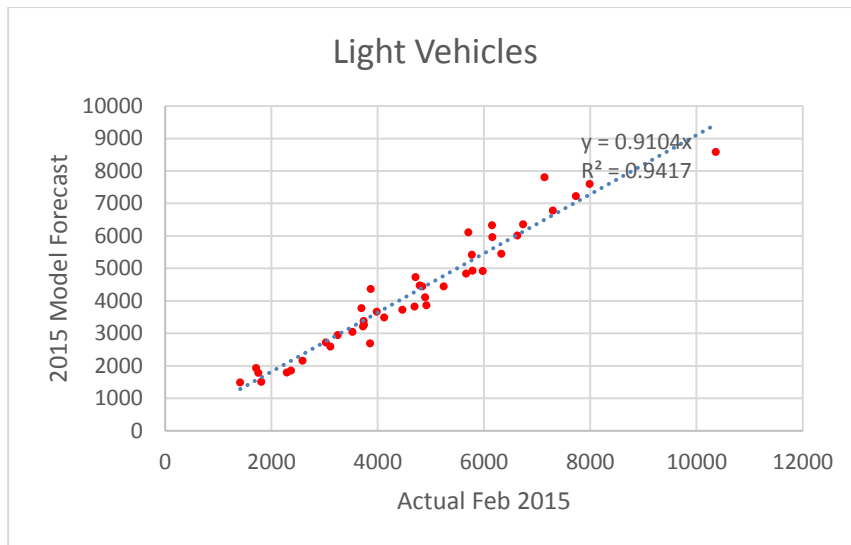


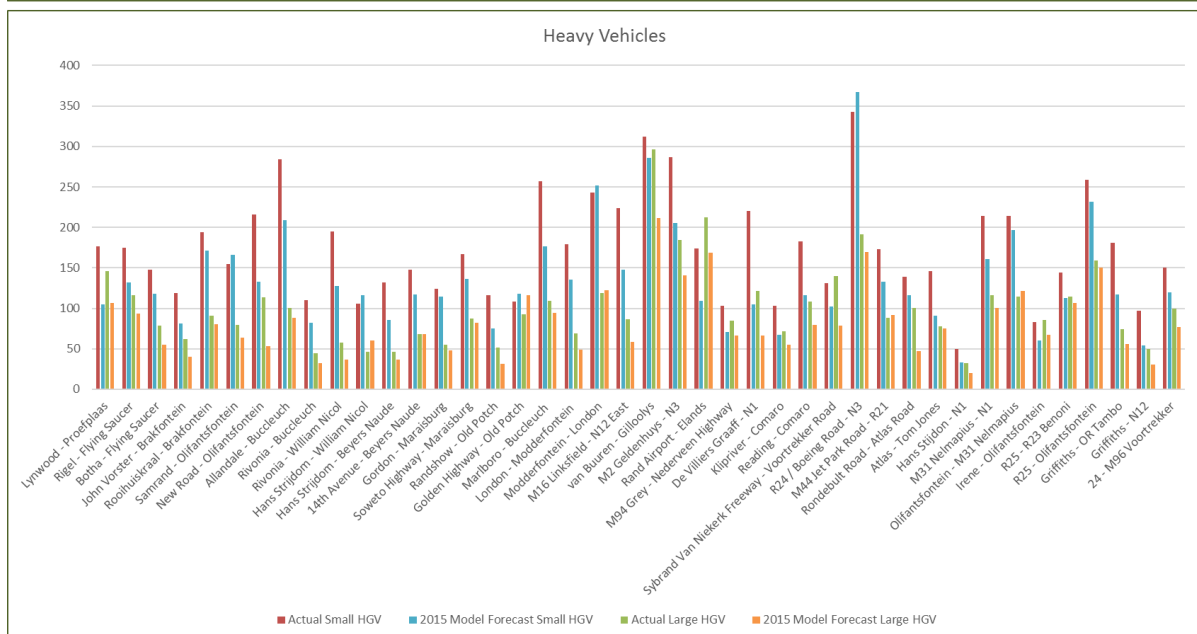
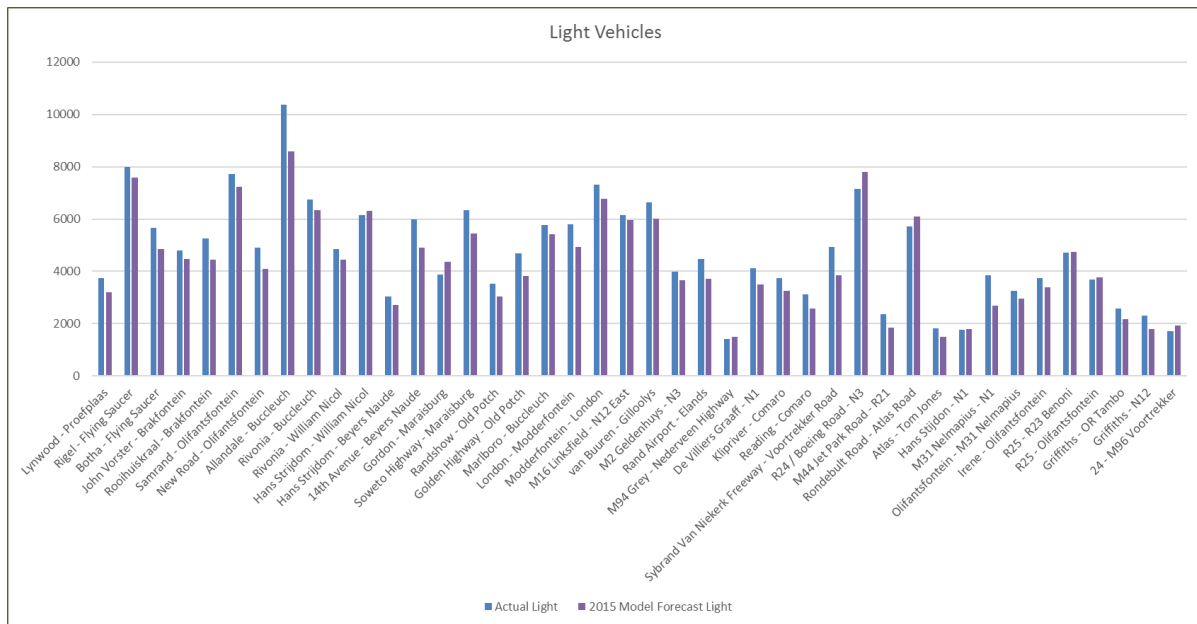
Note: The SL matrices were derived from a loaded network with all routes loaded with traffic, whereas the G2G matrix assignment is on an empty network with only one route loaded. The formula must be adjusted to only consider specific routes to allow or limit the use of multiple routes.

Appendix H: Gantry / Model Comparisons

Traffic Volumes at the Gantry Locations

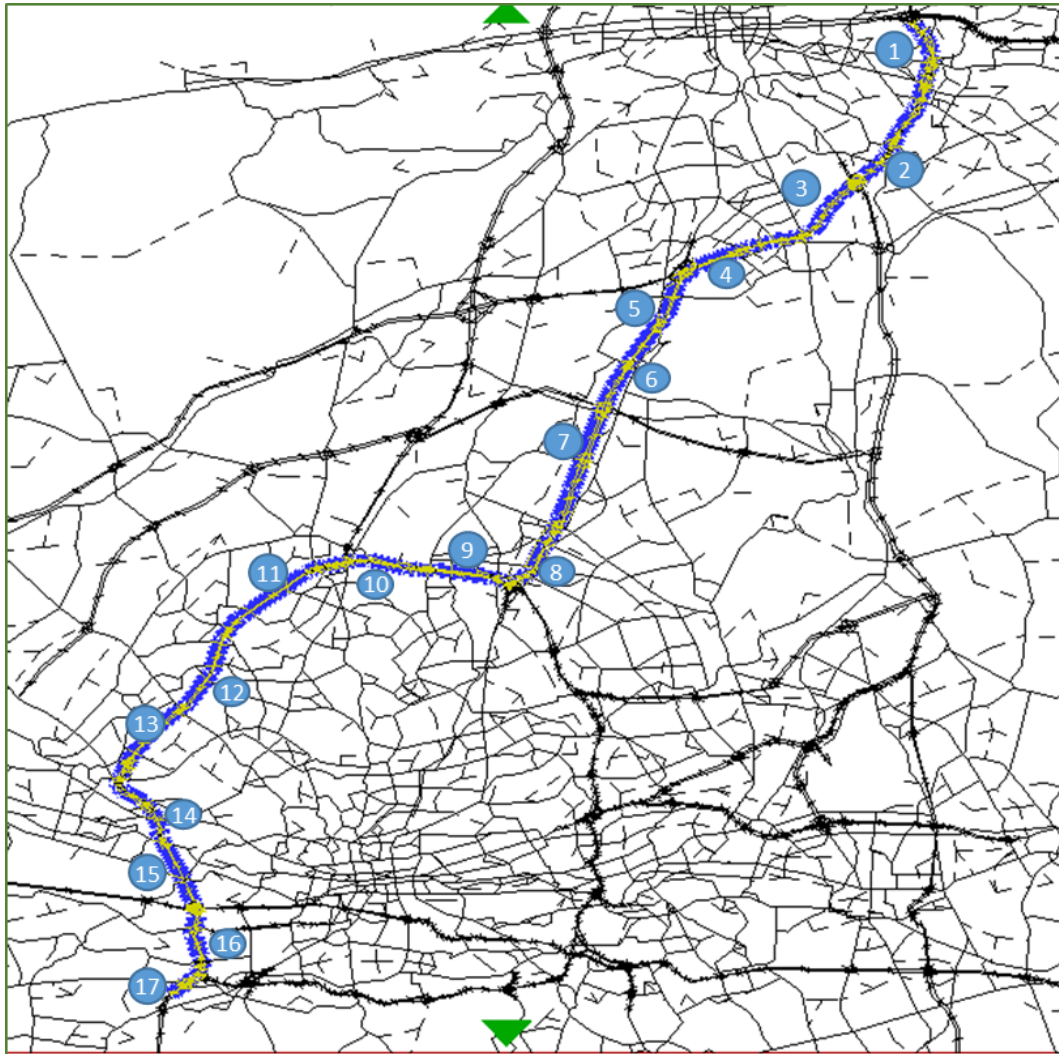
Location	Gantry	Actual			2015 Model Forecast			Lt %diff	Hgv %diff
		Light	Small HGV	Large HGV	Light	Small HGV	Large HGV		
Lynwood - Proefplaas	1	3728	176	146	3212	104	106	-14%	-35%
Rigel - Flying Saucer	2	7992	175	116	7595	132	93	-5%	-23%
Botha - Flying Saucer	3	5665	148	79	4841	118	55	-15%	-24%
John Vorster - Brakfontein	4	4795	119	62	4472	81	40	-7%	-33%
Rooihuiskraal - Brakfontein	5	5244	194	91	4443	171	80	-15%	-12%
Samrand - Olifantsfontein	6	7731	155	79	7221	166	64	-7%	-2%
New Road - Olifantsfontein	7	4897	216	113	4105	133	53	-16%	-44%
Allandale - Buccleuch	8	10366	284	100	8584	209	88	-17%	-23%
Rivonia - Buccleuch	9	6738	110	45	6352	82	32	-6%	-26%
Rivonia - William Nicol	10	4841	195	57	4440	127	37	-8%	-35%
Hans Strijdom - William Nicol	11	6157	105	46	6324	116	60	3%	16%
Hans Strijdom - Beyers Naude	12	3028	132	46	2718	85	37	-10%	-32%
14th Avenue - Beyers Naude	13	5977	148	68	4915	117	68	-18%	-14%
Gordon - Maraisburg	14	3871	124	55	4367	115	48	13%	-9%
Soweto Highway - Maraisburg	15	6327	167	87	5452	136	82	-14%	-14%
Randshow - Old Potch	16	3531	116	52	3043	75	32	-14%	-37%
Golden Highway - Old Potch	17	4700	108	93	3824	118	116	-19%	17%
Marlboro - Buccleuch	18	5776	257	109	5422	176	94	-6%	-26%
London - Modderfontein	19	5785	179	69	4928	135	49	-15%	-26%
Modderfontein - London	20	7301	243	119	6779	252	122	-7%	3%
M16 Linksfield - N12 East	21	6161	223	87	5962	148	58	-3%	-33%
van Buuren - Gilloolys	22	6630	312	296	6008	286	212	-9%	-18%
M2 Geldenhuys - N3	23	3986	287	185	3665	205	140	-8%	-27%
Rand Airport - Elands	24	4466	174	213	3723	109	169	-17%	-28%
M94 Grey - Nederveen Highway	25	1411	103	85	1487	71	66	5%	-27%
De Villiers Graaff - N1	28	4124	221	121	3485	105	66	-15%	-50%
klipriver - comaro	29	3743	103	72	3262	67	55	-13%	-30%
Reading - Comaro	30	3113	183	108	2586	116	79	-17%	-33%
Sybrand Van Niekerk Freeway - Voortrekke	31	4918	131	140	3856	102	78	-22%	-34%
R24 / Boeing Road - N3	32	7143	343	191	7805	367	169	9%	0%
M44 Jet Park Road - 21	33	2371	173	88	1851	133	92	-22%	-14%
Rondebult Road - Atlas Road	34	5711	139	101	6107	116	47	7%	-32%
atlas - Tom Jones	35	1810	146	78	1502	90	75	-17%	-26%
Hans Stijdon - N1	37	1758	50	32	1779	33	20	1%	-36%
M31 Nelmapius - N1	38	3860	215	117	2690	161	100	-30%	-21%
Olifantsfontein - M31 Nelmapius	39	3249	214	114	2947	197	122	-9%	-3%
irene - Olifantsfontein	40	3738	83	86	3382	60	68	-10%	-24%
25 - R23 Benoni	41	4713	145	114	4730	112	106	0%	-15%
25 - Olifantsfontein	42	3698	259	159	3774	231	151	2%	-9%
Griffiths - OR Tambo	43	2588	181	74	2159	117	56	-17%	-32%
Griffiths - N12	44	2293	97	50	1792	54	30	-22%	-42%
24 - M96 Voortrekker	45	1713	150	99	1930	120	77	13%	-21%

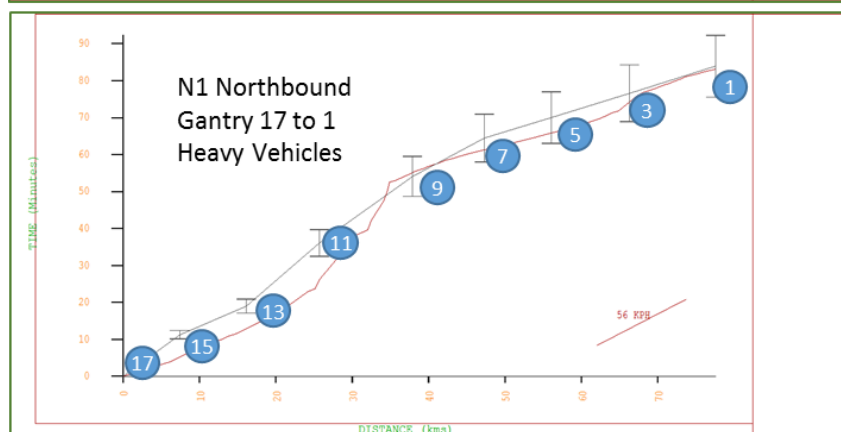
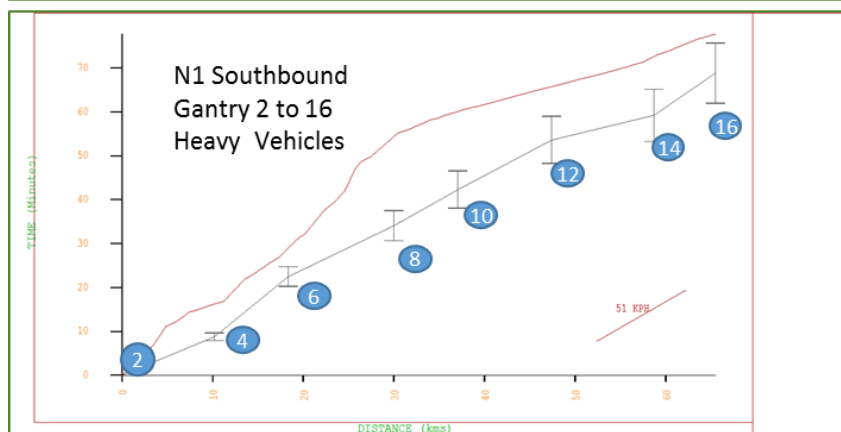
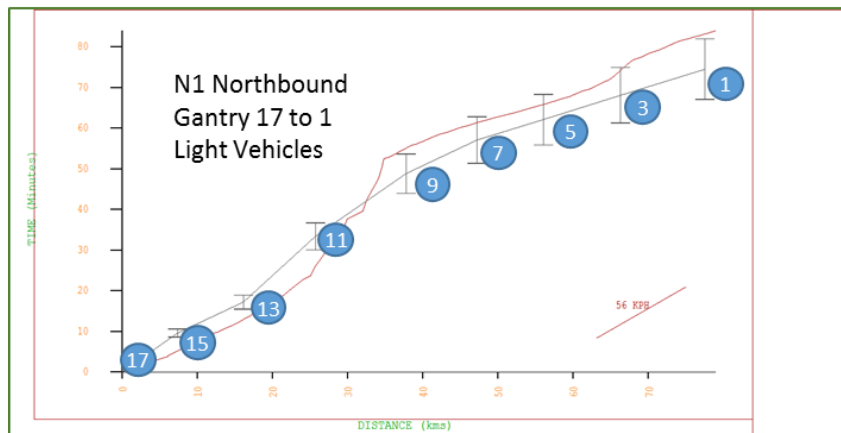
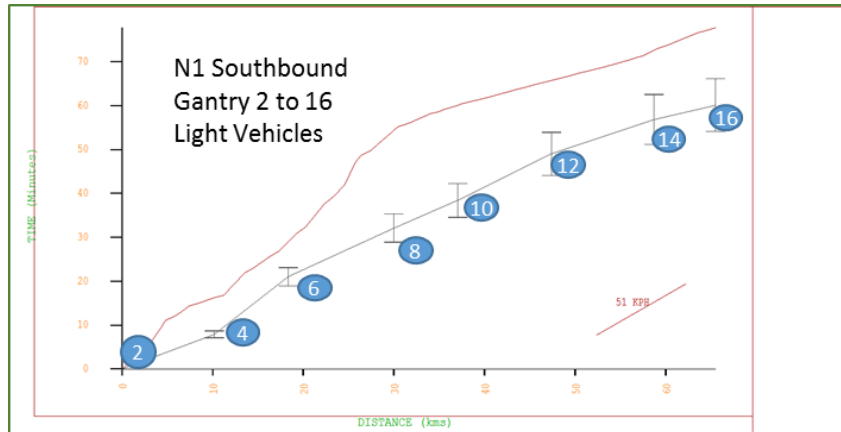




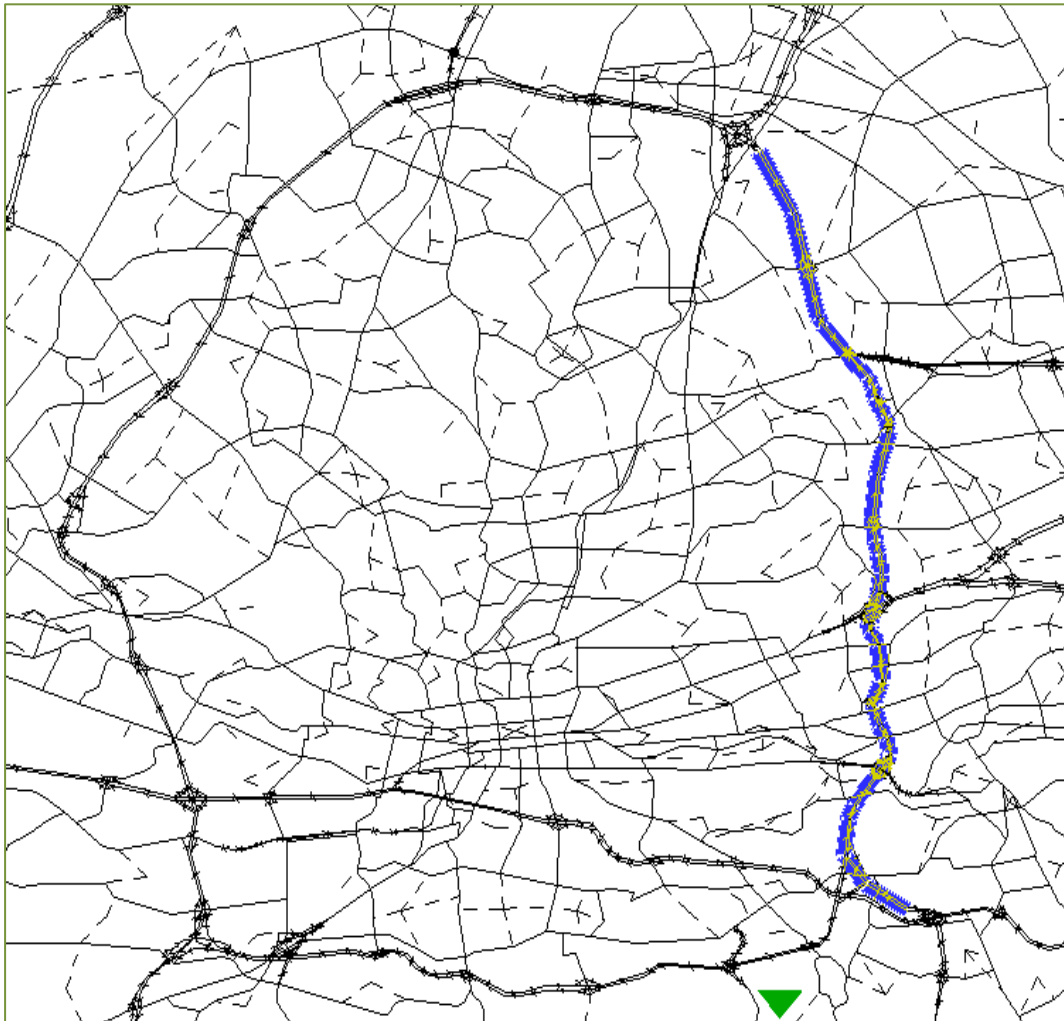
Journey Time Comparison

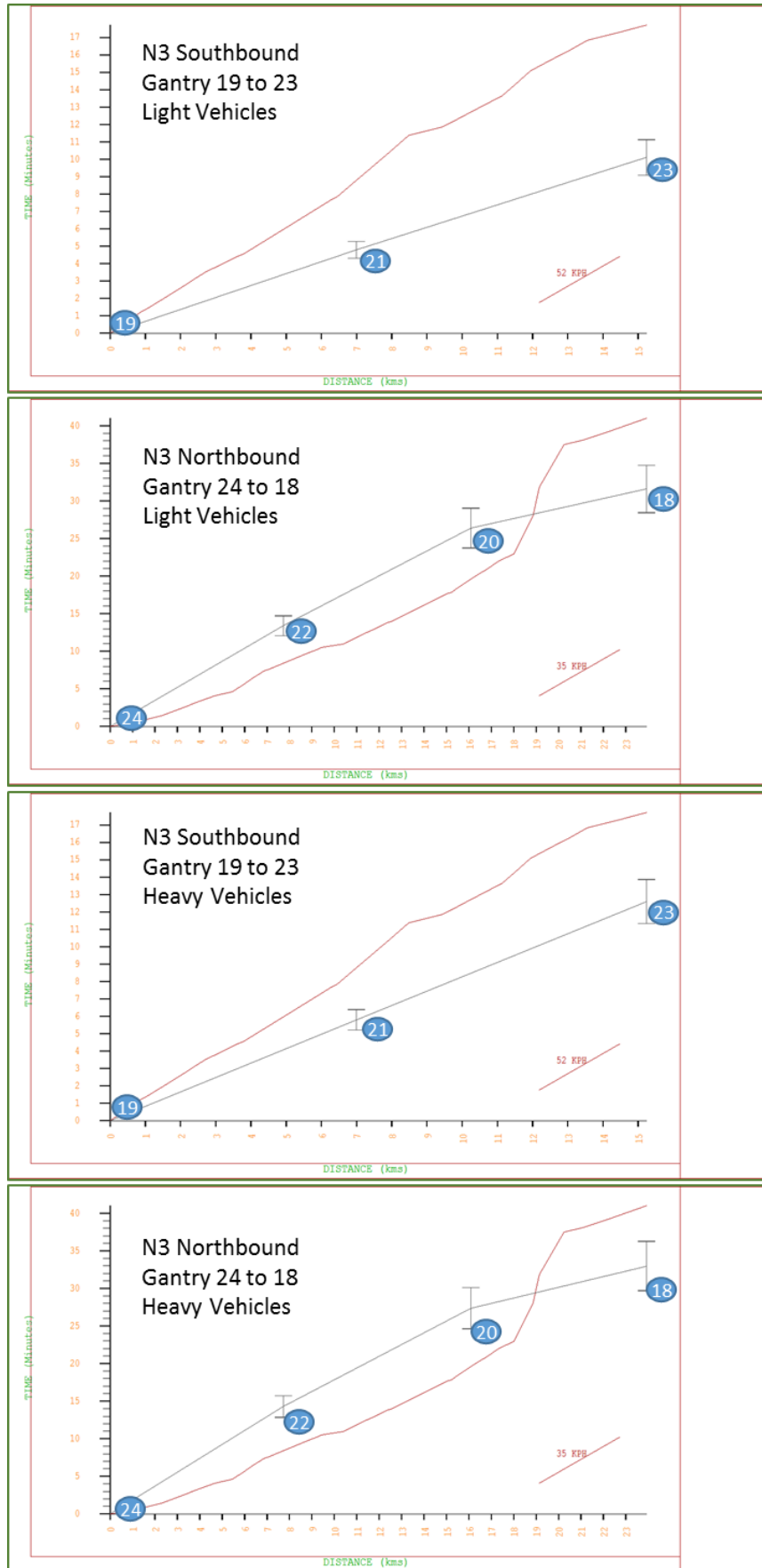
Route: N1



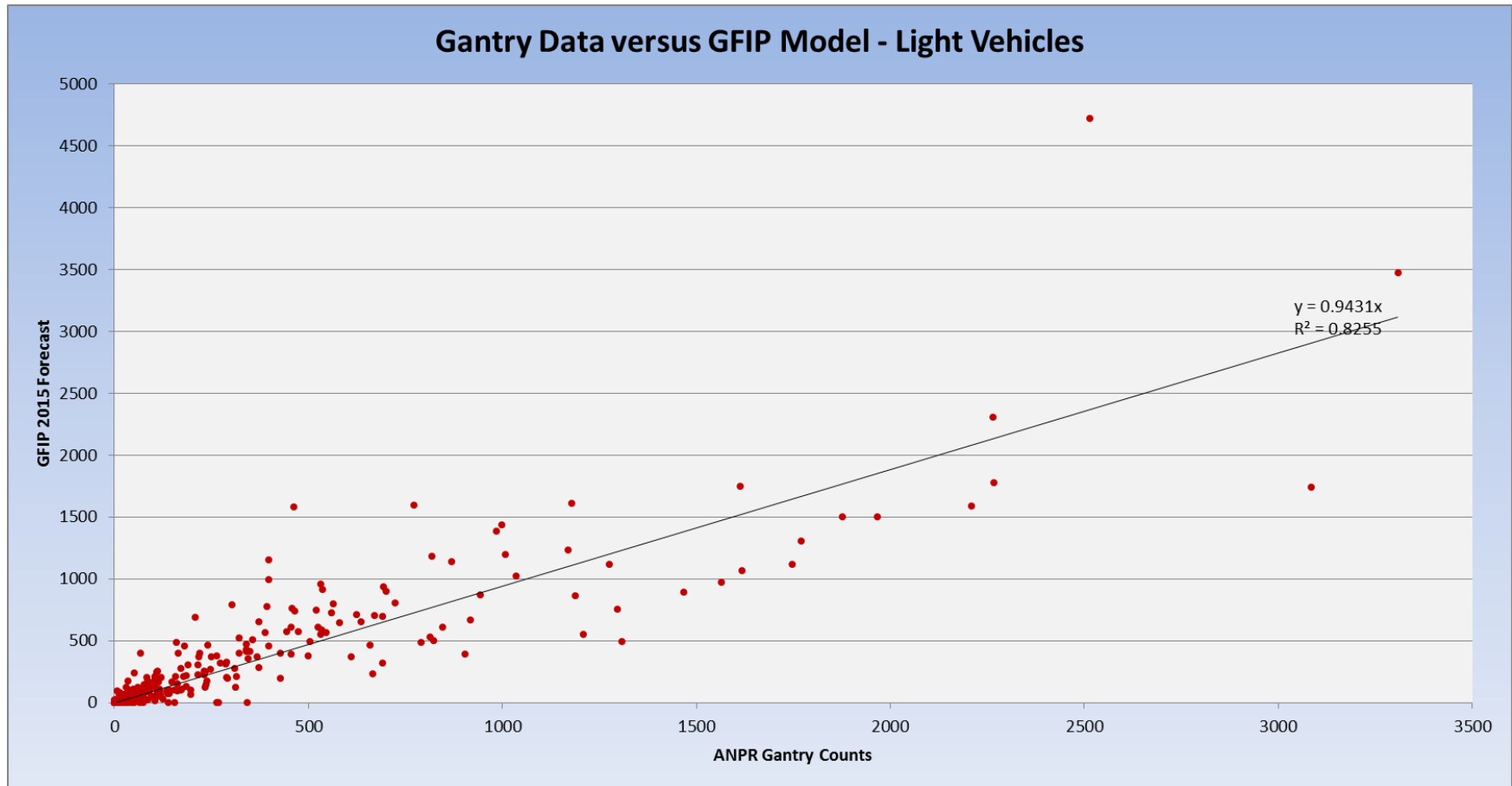


Route: N3





Gantry Counts to 2015 Model Output: Light Vehicles



IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

To	1001		1002		1003		1004		1005		1006		1007		1008		1009		1010		1011		1012		1013		1014		1015	
From	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model
1001	1467	896	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1002			3308	3478	0	0	1769	1309	3	0	536	917	0	0	531	958	0	0	85	22	0	0	38	11	0	0	9	1	0	
1003	869	1139	0	0	2267	1777	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1004	0	0	0	0	0	0	660	466	3	0	165	399	0	0	160	484	0	0	25	22	0	0	11	3	0	3	2	0	0	
1005	214	225	0	0	500	378	7	0	1276	1120	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1006	0	0	0	0	0	0	0	0	0	0	2208	1589	0	0	2264	2307	0	0	311	123	0	0	135	74	0	42	73	0	0	
1007	234	249	0	0	455	393	0	0	1169	1235	0	0	822	500	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1008	0	0	0	0	0	0	0	0	0	0	0	0	0	3084	1742	0	0	723	807	0	0	273	321	0	65	88	0	0	0	
1009	60	50	0	0	133	92	0	0	238	172	0	0	184	215	4	0	1618	1069	0	0	0	0	0	0	0	0	0	0	0	
1010	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	691	700	0	0	321	398	0	53	92	0	0	0	
1011	28	58	0	0	54	60	0	0	103	129	0	0	113	168	2	0	819	1180	5	0	1748	1118	0	0	0	0	0	0	0	
1012	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	692	322	0	341	470	0	0	0	
1013	13	48	0	0	21	34	0	0	46	104	0	0	53	57	0	0	357	511	0	0	814	529	5	1208	555	0	0	0	0	
1014	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	772	1598	0	0	0	0	
1015	6	8	0	0	10	6	0	0	36	53	0	0	22	40	0	0	76	143	0	0	302	791	0	789	490	7	1035	1022	0	
1016	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1017	0	0	0	0	12	25	0	0	33	47	0	0	21	37	0	0	83	88	0	0	241	466	0	664	236	0	903	396	0	
1018	15	17	0	0	24	13	0	0	69	88	0	0	72	62	0	0	0	0	264	380	0	0	88	144	0	9	8	0	0	
1019	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1020	21	5	0	0	28	1	0	0	117	65	0	0	141	76	0	0	0	0	525	607	0	0	162	97	0	7	0	0	0	
1021	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1022	34	62	0	0	14	20	0	0	74	43	0	0	74	57	0	0	0	0	108	111	0	0	7	0	0	0	13	0	0	
1023	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	4	0	51	0	0	0	
1024	18	14	0	0	8	0	0	0	54	19	0	0	61	43	0	0	0	0	107	56	0	0	3	6	89	126	0	184	133	
1025	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1026	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1027	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1028	0	0	0	0	0	0	0	0	2	6	0	0	1	6	0	0	6	1	0	0	109	245	0	372	287	0	473	573	0	
1029	3	0	0	0	1	0	0	0	5	2	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1030	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	33	46	0	197	101	0	291	198	0	
1031	23	29	0	0	9	4	0	0	42	48	0	0	33	68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1032	1	0	0	0	1	0	0	0	8	0	0	0	29	4	0	0	0	0	230	255	0	0	56	85	6	97	0	51	51	
1033	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1035	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1037	0	0	0	0	110	80	0	0	0	34	67	0	0	31	82	0	0	5	1	0	0	1	1	0	0	0	0	0	0	
1038	85	140	3	0	125	28	0	0	34	20	0	0	36	24	0	0	6	6	0	0	2	0	0	0	0	0	0	0	0	
1039	66	92	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1040	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1041	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1042	288	328	2	0	75	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1043	77	108	0	0	10	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1044	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1045	70	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

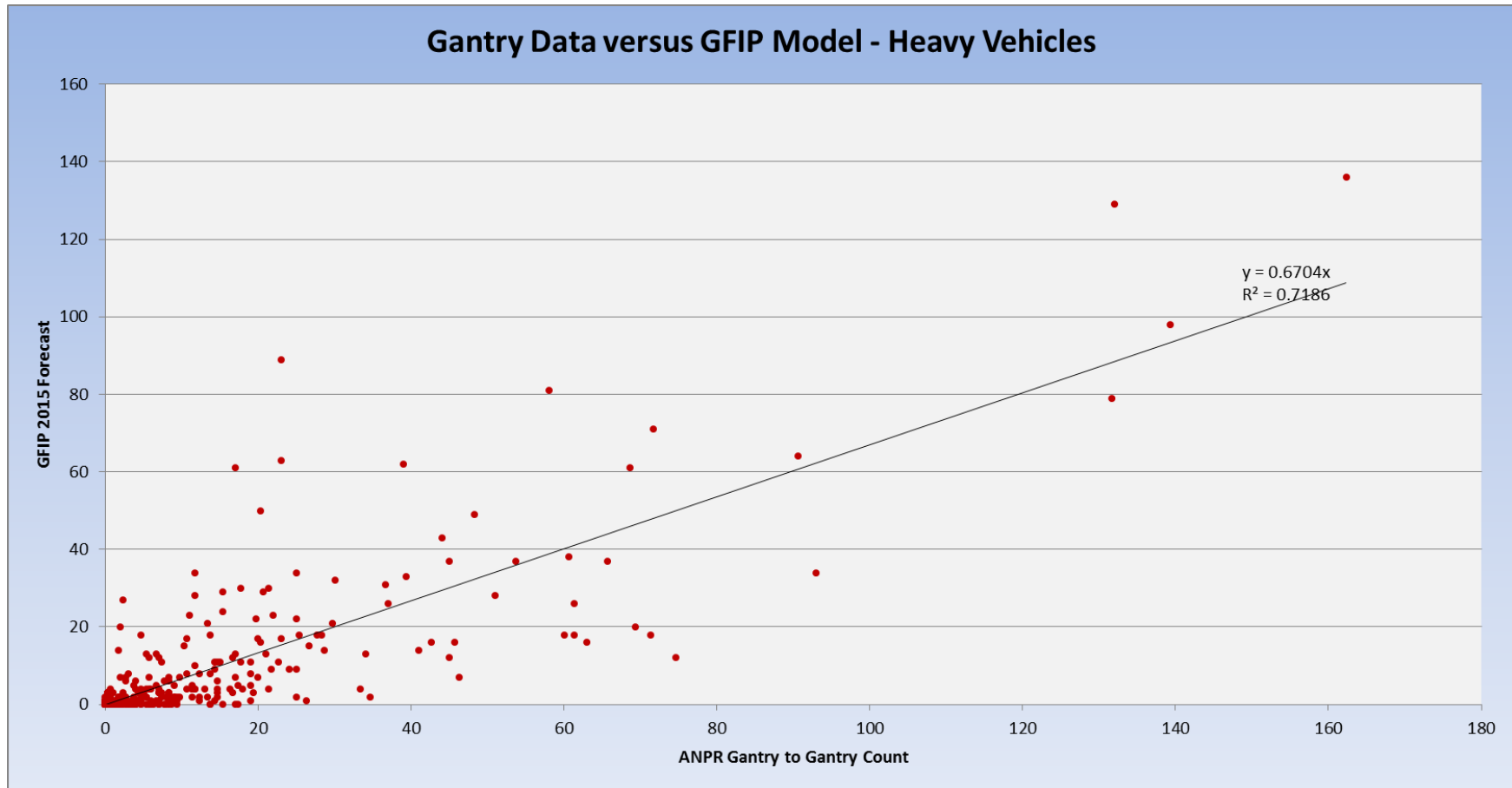
IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

To	1016		1017		1018		1019		1020		1021		1022		1023		1024		1025		1028		1029		1030	
From	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model
1001	0		0		0		0		0		0		0		0		0		0		0		0		0	
1002	19	2	0		0		30	121	1		15	5	0		57	35	0		0		5	3	0		9	1
1003	0		0		0		0		0		0		0		4		0		1		0		0		0	
1004	4	1	0		0		16	38	0		14	22	0		9	4	0		2	0	0		0		0	0
1005	0		0		0		0		0		0		0		0		0		0		0		0		0	
1006	46	53	0		1		216	302	0		221	399	0		106	60	0		40	14	2	4	0		10	3
1007	0		0		0		0		0		0		0		0		0		0		0		0		0	
1008	52	99	0		13		456	612	0		579	644	0		156	101	0		48	30	5	8	0		21	5
1009	0		0		0		545	570	0		700	897	0		117	96	0		37	43	0	0	0		3	0
1010	49	112	0		0		0		0		0		0		0		0		0		0		1	0	0	
1011	0		0		0		286	316	0		233	123	0		8	0	0		5	0	0		0		0	
1012	339	417	0		0		0		0		0		3	0	0		0		10	8	0		82	93	0	
1013	0		0		0		104	18	0		43	0	0		0	4	0		0	0	0		0		0	
1014	942	872	0		0		0		0	0	0		64	6	0		0		53	6	0		314	209	0	
1015	18		0		0		2	0	0		0		0		0		0		0		0		0		0	
1016	531	549	0		0		0		1	0	0		32	1	0		0		24	7	0		149	170	0	
1017	0		1877	1502	2	0	2	0	11	1	0		141	102	0		0		0		0		208	692	0	
1018	1	0	0		367	368	0		0		0		0		0		0		0		0		0		0	
1019	0		0		0		397	1157	0		462	1583	0		104	213	0		35	89	0		0		11	9
1020	1	0	0		999	1435	0		917	671	0		0		0		0		0		0		0		0	
1021	0		0		0		0		0		1613	1752	0		398	460	0		101	174	60	60	0		78	32
1022	0		0		371	652	0		339	426	0		1563	972	0		0		0		0		0		0	
1023	0		0		0		0		0		0		0		693	937	0		444	575	95	50	0		116	51
1024	0		0		344	354	0		289	205	0		670	703	0		1965	1502	0		216	368	0		236	147
1025	0		0		0		0		0		0		0		0		0		321	520	0		0		0	
1026	0		0		0		0		0		0		0		0		0		0		0		0		0	
1027	0		0		0		0		0		0		0		0		0		0		0		0		0	
1028	0		0		0		0		0		0		0		0		0		0		1296	758	0		0	
1029	0		0		16	0	0		18	5	0		174	110	0		0		57	15	0		428	398	0	
1030	0		0		0		0		0		0		0		0		0		0		250	373	0		636	655
1031	0		0		162	155	0		113	99	0		457	763	0		0		33	6	0		0		0	
1032	0		0		393	775	0		397	998	0		0		388	564	0		87	96	111	189	0		91	96
1033	0		0		0		0		0		0		0		0		0		0		0		0		0	
1035	0		0		0		0		0		0		0		0		0		0		0		0		0	
1037	1	1	0		0		2	25	0		1	16	0		14	18	0		4	8	0		0		0	
1038	0		0		0		1	0	0		0		0		0		0		0		0		0		0	
1039	0		0		0		0		0		0		0		0		0		0		0		0		0	
1040	0		0		0		0		0		0		0		18	37	0		5	3	0		0		0	
1041	0		0		1	0	0		2	4	0		0		23	64	0		7	4	0		0		0	
1042	0		0		0		0		0		0		0		0		0		0		0		0		0	
1043	0		0		0		0		0		0		0		0		0		0		0		0		0	
1044	0		0		0		0		0		0		0		13	15	0		2	2	0		0		0	
1045	0		0		0		0		0		0		0		0		0		0		0		0		0	

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

To	1031		1032		1033		1035		1037		1038		1039		1040		1041		1042		1043		1044		1045	
From	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model
1001	0		0		0		0		0		0		0		0		0		0		0		0		0	
1002	0		0		0		23		2		15		0		248	271	846	610	0		1		171	100	0	
1003	0		14		0		11	4	3		4		0		63	78	341		0		0		61	77	0	
1004	0		0		0		0		0		0		0		0		0		0		0		0		0	
1005	0		0		0		1		0		1		0		3	0	35	0	0		0		0		7	
1006	0		0		5	0	0		0		0		0	0	0		0		0		0		0		0	0
1007	0		0		0		0		0		0		0		0		0		0		0		0		0	
1008	0	0	0		43	24	16	10	0		0		0		0		0		0		0		0		0	0
1009	0		0		83	202	27	91	0		0		0		0		0		0		1		1		47	13
1010	0	0	0		0		0		0		0		0		0		0		0		0		0		0	
1011	0		0		38	97	12	46	0		0		0		0		0		0		0		0		25	9
1012	51	59	0		0		0		0		0		0		0		0		0		0		0		0	
1013	0		0		7	0	2		0		0		0		0		0		0		0		0		0	
1014	428	195	0		0		0		0		0	0		0		0		0		0		0		0	5	1
1015	0		0		0		0		0		0		0		0		0		0		0		0		0	
1016	172	275	0		7	4	1	1	0		0		0		0		0		0		0		0		5	1
1017	350	413	0		12	70	3		0		1	21	0	14		0		1	12	0		0		20	57	
1018	0		0		0		0		0		0		0		0		0		0		0		0		0	
1019	0		0		67	401	21	195	0		0		0		0		0		0		0		0		37	32
1020	0		0		0		0		0		0		0		0		0		0		0		0		0	
1021	0		0		86	173	43	93	0		0		0		0		0		0		0		0		0	
1022	0		0		119	205	38	37	0		54	58	12	25	0		0		23	45	3		4		190	303
1023	4		0		0		0		0		0		0		0		0		0		0		0		0	
1024	0		0		15	0	0		0		27	14	5	12	0		0		9	21	0		0		60	123
1025	0		0		0		0		0		0		0		0		0		0		0		0		0	
1026	0		0		0		0		0		0		0		0		0		0		0		0		0	
1027	0		0		0		0		0		0		0		0		0		0		0		0		0	
1028	0		0		0		0		0		0		0		0		0		0		0		0		0	
1029	624	709	0		15	78	4	8	0		5	2	2	3	0		0		2	1	0		0		20	52
1030	0		0		0		0		0		0		0		0		0		0		0		0		0	
1031	1307	491	0		34	175	8	22	0		31	65	6	30	0		0		10	19	0		0		50	241
1032	1		2513	4719	0		0		0		0		0		0		0		0		0		0		0	
1033	0		0		563	796	284	433	0		64	0	7	0	0		0		27	1	156	3	0		0	
1035	0		0		0		265	362	0		0		0		0		0		0		0		0		0	
1037	0		49	26	0		0		559	725	0		0		158	208	503	496	0		0		86	89	0	
1038	0		0		0		0		0		610	369	0		0		0		0		0		0		0	
1039	0		0		0		0		0		309	280	21	66	0		0		0		0		0		0	
1040	0		72	54	0		15	2	0		0		0		115	111	465	742	0		0		89	120	0	
1041	0		117	87	0		17	10	0		0		0		0		1178	1609	0		0		178	211	0	
1042	0		0		0		0		1		1189	863	180	459	0		0		520	745	0		0		0	
1043	0		0		0		0		0		231	225	38	102	0		0		111	256	984	1386	25		1	
1044	0		75	97	0		144	197	0		1		0		0		0		0		22		1007	1197	0	
1045	0		0		0		0		0		197	63	30	39	0		0		64	94	0		0		534	592

Gantry Counts to 2015 Model Output: Heavy Vehicles



IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

To	1001		1002		1003		1004		1005		1006		1007		1008		1009		1010		1011		1012		1013		1014		1015	
From	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model
1001	51	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1002	0	0	66	37	0	0	69	20	0	0	11	5	0	0	10	15	0	0	2	0	0	0	0	0	0	0	1	0	0	
1003	41	14	0	0	37	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1004	0	0	0	0	0	0	25	2	0	0	2	2	0	0	2	7	0	0	0	0	0	0	0	0	0	0	0	0	0	
1005	23	17	0	0	15	29	0	0	58	81	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1006	0	0	0	0	0	0	0	0	0	0	45	37	0	0	48	49	0	0	7	3	0	0	5	2	0	0	4	5	0	
1007	17	0	0	0	12	1	0	0	46	7	0	0	60	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1008	0	0	0	0	0	0	0	0	0	0	0	0	0	93	34	0	0	29	14	0	0	13	21	0	0	7	11	0	0	
1009	1	0	0	0	0	0	0	0	2	0	0	0	5	1	0	0	19	3	0	0	0	0	0	0	0	0	0	0	0	
1010	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	9	0	0	9	5	0	0	14	0	0	0	
1011	2	0	0	0	1	1	0	0	2	1	0	0	4	1	0	0	15	11	0	0	19	8	0	0	0	0	0	0	0	
1012	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	8	0	0	20	7	0	0	
1013	2	2	0	0	1	4	0	0	3	7	0	0	4	2	0	0	7	13	0	0	11	23	0	0	46	16	0	0	0	
1014	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	22	0	0	
1015	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	3	2	0	0	35	2	0	0	34	13
1016	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1017	0	0	0	0	0	0	0	0	2	1	0	0	2	0	0	0	2	2	0	0	2	27	0	0	20	17	0	0	17	61
1018	6	0	0	0	7	1	0	0	8	2	0	0	15	2	0	0	0	0	17	7	0	0	8	7	0	0	1	1	0	0
1019	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1020	2	0	0	0	0	0	0	0	5	4	0	0	10	7	0	0	0	0	15	3	0	0	5	2	0	0	0	0	0	0
1021	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1022	17	0	0	0	1	0	0	0	11	4	0	0	17	13	0	0	0	0	8	2	0	0	1	0	0	0	0	0	0	0
1023	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	5	0
1024	13	2	0	0	0	0	0	0	8	3	0	0	14	8	0	0	0	0	4	1	0	0	0	0	6	0	0	0	14	0
1025	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1026	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1027	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1028	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	9	2	0	0	17	3
1029	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1030	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	13	4	0	0	24	9
1031	6	1	0	0	0	0	0	0	3	1	0	0	5	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1032	0	0	0	0	0	0	0	0	1	0	0	0	4	0	0	0	0	0	21	13	0	0	9	2	1	1	0	0	9	1
1033	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1035	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1037	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1038	4	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1039	15	24	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1040	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1041	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1042	43	16	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1043	17	5	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1044	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1045	18	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

To	1016		1017		1018		1019		1020		1021		1022		1023		1024		1025		1028		1029		1030	
From	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model
1001	0		0		0		0		0		0		0		0		0		0		0		0		0	
1002	2	1	0		0		2	2	0		1	0	0		6	0	0		0		0	0	0		0	0
1003	0		0		0		0		0		0		0		0		0		0		0		0		0	
1004	0	0	0		0		0	1	0		0	2	0		0	0	0		0	0	0		0		0	0
1005	0		0		0		0		0		0		0		0		0		0		0		0		0	
1006	6	4	0		0		8	6	0		15	11	0		11	2	0		6	4	0	0	0		1	0
1007	0		0		0		0		0		0		0		0		0		0		0		0		0	
1008	5	4	0		0		15	6	0		19	5	0		14	1	0		7	3	0	0	0		1	0
1009	0		0		0		7	1	0		7	4	0		1	0	0		1	1	0	0	0		0	
1010	0	1	0		0		0		0		0		0		0		0		0		0		0	0	0	
1011	0		0		0		9	2	0		7	2	0		0	0	0		0	0	0		0		0	
1012	14	9	0		0		0		0		0		0	0	0		0		1	0	0		5	0	0	
1013	0		0		0		2	0	0		0	0	0		0	0	0		0	0	0		0		0	
1014	23	11	0		0		0		0	0	0		1	0	0		0		4	0	0		8	1	0	
1015	0		0		0		0	0	0		0		0		0		0		0		0		0		0	
1016	33	4	0		0		0		0		0		1	0	0		0		4	0	0		9	0	0	
1017	0		63	16	0	0	0	0	0	0		0		0	0	0		0		0		12	4	0		
1018	0	0			27	15	0		0		0		0		0		0		0		0		0		0	
1019	0		0		0		19	11	0		21	29	0		7	5	0		3	8	0		0		0	0
1020	0	0			17	12	0		22	9	0		0		0		0		0		0		0		0	
1021	0		0		0		0		0		61	26	0		21	4	0		7	12	8	0	0		3	0
1022	0		0		22	23	0		20	22	0		124	80	0		0		0		0		0		0	
1023	0		0		0		0		0		0		0		61	18	0		54	37	19	1	0		4	0
1024	0		0		12	34	0		11	17	0		68	74	0		139	98	0		26	1	0		7	0
1025	0		0		0		0		0		0		0		0		0		18	11	0		0		0	
1026	0		0		0		0		0		0		0		0		0		0		0		0		0	
1027	0		0		0		0		0		0		0		0		0		0		0		0		0	
1028	0		0		0		0		0		0		0		0		0		0		75	12	0		0	
1029	0		0		0	0		0	0		0		3	1	0		0		4	0	0		12	2	0	
1030	0		0		0		0		0		0		0		0		0		0		39	62	0		20	50
1031	0		0		4	3	0		3	1	0		25	21	0		0		2	0	0		0		0	
1032	0		0		18	30	0		23	89	0		0		72	71	0		28	18	28	18	0		7	3
1033	0		0		0		0		0		0		0		0		0		0		0		0		0	
1035	0		0		0		0		0		0		0		0		0		0		0		0		0	
1037	0	0			0		0	0	0		0	0		1	0	0		0	2	0		0		0		0
1038	0		0		0		0	0	0		0		0		0		0		0		0		0		0	
1039	0		0		0		0		0		0		0		0		0		0		0		0		0	
1040	0		0		0		0		0		0		0		0		0		0		0		0		0	
1041	0		0		0	0		0	0	1		0		0		4	1	0		1	0	1		0		0
1042	0		0		0		0		0		0		0		0		0		0		0		0		0	
1043	0		0		0		0		0		0		0		0		0		0		0		0		0	
1044	0		0		0		0		0		0		0		1	0	0		0		0		0		0	
1045	0		0		0		0		0		0		0		0		0		0		0		0		0	

IMPROVING TRAFFIC MODELS USING LARGE-SCALE AUTOMATIC NUMBER PLATE RECOGNITION (ANPR) DATA

To	1031		1032		1033		1035		1037		1038		1039		1040		1041		1042		1043		1044		1045		
From	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	Gantry	Model	
1001	0		0		0		0		0		0		0		0		0		0		0		0		0		
1002	0		0		0		0		0		0		0		20	16	45	12	0		0		0		9	1	
1003	0		0		0		0		0		0		0		2	0	7		0		0		1	0	0		
1004	0		0		0		0		0		0		0		0		0		0		0		0		0		
1005	0		0		0		0		0		0		0		0	0	0	0	0		0		0		0		
1006	0		0		0	0	0		0		1		0		0		0		0		0		0		0		
1007	0		0		0		0		0		0		0		0		0		0		0		0		0		
1008	0	0	0		2	0	0	0		0		0		0		0		0		0		0		0		0	
1009	0		0		0	3	1	1		0		0		0		0		0		0		0		0		1	0
1010	0	0	0		0		0		0		0		0		0		0		0		0		0		0		
1011	0		0		0	3	0	2		0		0		0		0		0		0		0		0		1	0
1012	2	1	0		0		0		0		0		0		0		0		0		0		0		0		
1013	0		0		0	0	0		0		0		0		0		0		0		0		0		0		
1014	6	12	0		0		0		0		0	0		0		0		0		0		0		0		0	0
1015	0		0		0		0		0		0		0		0		0		0		0		0		0		
1016	8	3	0		0	0	0	0		0		0		0		0		0		0		0		0		0	0
1017	21	30	0		1	0	0		0		0	0		0	0		0		1	0	0		0		2	0	
1018	0		0		0		0		0		0		0		0		0		0		0		0		0		
1019	0		0		2	20	1	8		0		0		0		0		0		0		0		0		1	0
1020	0		0		0		0		0		0		0		0		0		0		0		0		0		
1021	0		0		5	13	2	6		0		0		0		0		0		0		0		0		0	
1022	0		0		12	4	8	4		0	10	2		1	3		0		11	8	0		0		30	21	
1023	0		0		0		0		0		0		0		0		0		0		0		0		0		
1024	0		0		2	0	0	0		0	3	0		1	0		0		5	3	0		0		11	4	
1025	0		0		0		0		0		0		0		0		0		0		0		0		0		
1026	0		0		0		0		0		0		0		0		0		0		0		0		0		
1027	0		0		0		0		0		0		0		0		0		0		0		0		0		
1028	0		0		0		0		0		0		0		0		0		0		0		0		0		
1029	12	28	0		0	0	0	0		0		0	0		0		0		0	0	0		0		0	0	
1030	0		0		0		0		0		0		0		0		0		0		0		0		0		
1031	71	18	0		2	14	1	13		0	1	1		0	1		0		3	2	0		0		6	7	
1032	0		162	136	0		0		0		0		0		0		0		0		0		0		0		
1033	0		0		132	129	19	78		0	9	0		1	0		0		8	0	15	0	0		0		
1035	0		0		0		17		0		0		0		0		0		0		0		0		0		
1037	0		0	3	0		0		12	10	0		0		11	4	16	4	0		0		3	0	0		
1038	0		0		0		0		0		15	4		0		0		0		0		0		0		0	
1039	0		0		0		0		0		30	32		4	4		0		0		0		0		0		
1040	0		0	0	0		0		0		0		0		4	0	3	0	0		0		0		0	0	
1041	0		8	6	0		2	0	0		0		0		0		91	64	0		0		14	11	0		
1042	0		0		0		0		0		39	33		5	18		0		61	38	0		0		0		
1043	0		1		0		0		0		25	18		3	6		0		25	34	69	61	0		0		
1044	0		4	6	0		9	11	0		0		0		0		0		0		0		44	43	0		
1045	0		0		0		0		0		23	63		2	3		0		14	18	0		0		37	31	