USES OF ANPR DATA IN TRAFFIC MANAGEMENT AND TRANSPORT MODELLING

A ROBINSON and A VAN NIEKERK*

Hatch Goba (Pty) Ltd, Private Bag X20, Gallo Manor

Tel: 011-239-5614; Email: robinsona@hatch.co.za

*South African National Roads Agency (SOC) Limited;

Northern Region, 38 Ida Street, Menlo Park, Pretoria, 0081

Tel: 012 4266226 Email: NiekerkA@nra.co.za

ABSTRACT

Automatic Number Plate Recognition (ANPR) technology provides the opportunity to collect accurate traffic data from various locations throughout a road network. The data, collected from a wide area, can produce detailed information on traffic operations including traffic counts, average travel speeds and reliable origin destination (OD) data. This data can also provide essential inputs into the building, calibration and validation of traffic models. It is the author's belief that this data is under-utilised and this paper explores the potential uses and true value of this data, including the generation of origin-destination (OD) information.

1 INTRODUCTION

The Gauteng Freeway Improvement Project's toll gantries incorporate Automatic Number Plate Recognition (ANPR) technology. These systems collect accurate traffic data recording vehicles with a unique identity derived from the number plates with the time and the vehicle type. We must point out that the vehicle classification system is not part of the ANPR systems but is from the vehicle profiling technology.

A scan of the internet indicates that the use of ANPR data is predominantly used to calculating travel times between successive camera locations, with very little information relating to the use of this data in developing Origin-Destination (OD) trip matrices. Friedrich et al. (2008) briefly considered ANPR data to derive an OD matrix from a single point using the city prefix on a number plate and direction of travel to determine ODs on a national road network. Other "Big Data" sets such as Bluetooth detection (Michau, Nantes & Chung, 2013) also pose issues with respect to "cleaning" the data, determining the mode of travel and relating the output to a model's zone system. Van Vuren (2011) evaluated wide area GPS data for deriving an OD matrix for a major urban area, with 7 million entries over a year. Compared to observed flows this sample was about 4% for light goods vehicles, 1% or heavies and up to 0.75% for cars.

We are of the opinion that the data from the GFIP gantries and similar data sources is under-utilised and detailed analysis of this data, combining location time-stamps and vehicle counts will produce accurate traffic data for use in traffic planning and day-to-day traffic management.

ANPR traffic data provides accurate and comprehensive data sets from wide spread locations. This paper describes the data obtainable from the Gauteng Freeway Improvement Project (GFIP) systems and considers the potential uses of the data with respect to input into day-to day traffic management initiatives and in the development of traffic and transportation models.

2 GANTRY DATA

This evaluation uses data obtained from the Gauteng Freeway Open Road Toll gantries during September 2012. Figure 1 depicts the locations of the gantries along the Gauteng freeways, which are spaced approximately ten kilometres apart in each direction and offset by approximately five kilometres in each opposite direction. The gantry equipment captures images of every vehicle that passes under it, records the number plate, time and gantry number. An overall database is compiled from the information from all gantries.

Two query runs on the database provided the information used in this analysis. The first provides the basis for deriving classified traffic counts at each gantry location and the second records the gantries that a vehicle passes under while travelling on the freeways within a specified time-period. This latter data set records the first and last gantry that a vehicle passes while on the freeway.

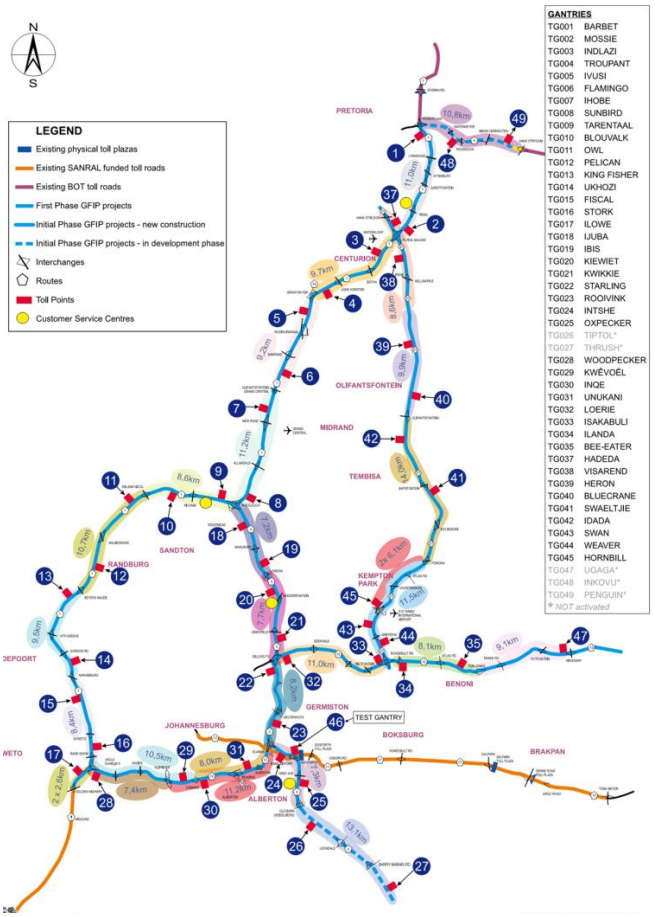

Figure 1: Gauteng Freeway Gantry Locations

Prior to receiving the data, the vehicles number plate information was replaced with a unique record number to ensure that any personal information was not divulged or be obtained.

3 TRAFFIC COUNTS

Table 1 provides a sample of the traffic count data obtained from the traffic database. Each data record comprises the gantry number, the date, the time (in 15-minute intervals, numbered consecutively in the table, i.e. 0=0min00sec to 14min59sec, 1=15min to 29min59sec etc.) and the number of vehicles passing under the gantry during the 15-minute period by toll classification. The vehicle classes are:

- Class A1 – motorcycles
- Class A2 – cars, minibus, sports utility vehicle (SUV)<2.5m high
- Class B – small heavy vehicle < 12m
- Class C – large heavy vehicle > 12m

Table 1: Sample of Traffic Count Data – Vehicles per 15 Minutes

GANTRY_ NUMBER	Date	HOUR	MIN15	A1	A2	B	C
1001	2012/09/01	5	3	3	305	13	44
1001	2012/09/01	6	0	1	334	15	35
1001	2012/09/01	6	1	3	377	21	43
1001	2012/09/01	6	2	1	441	19	33
1001	2012/09/01	6	3	5	558	21	33
1001	2012/09/01	7	0	1	590	29	33
1001	2012/09/01	7	1	11	679	19	40
1001	2012/09/01	7	2	3	800	30	42
1001	2012/09/01	7	3	9	777	18	18
1001	2012/09/01	8	0	4	763	38	17
1001	2012/09/01	8	1	7	779	39	21
1001	2012/09/01	8	2	48	848	24	27
1001	2012/09/01	8	3	26	859	26	25
1001	2012/09/01	9	0	10	838	30	12
1001	2012/09/01	9	1	8	900	24	21
1001	2012/09/01	9	2	8	865	32	16

Source: ETC Central Operations Centre

This data was summarised to provide hourly average traffic counts by vehicle class for each toll gantry. Table 2 includes the summarised traffic counts for toll gantry numbers 3 and 5 for an average weekday during September 2012. The use of this information in traffic model development is mentioned in Section 6.3 below.

Table 2: Summarised Traffic Counts

Gantry		1003					1005				
Vehicle Class		A1	A2	B	C	Total	A1	A2	B	C	Total
Time	00:00	1	199	14	18	232	1	240	20	25	286
	01:00	1	112	15	21	150	1	142	21	29	193
	02:00	0	88	15	21	125	1	106	22	28	156
	03:00	0	107	13	24	144	1	116	20	29	166
	04:00	1	291	25	27	344	2	256	37	42	336
	05:00	4	917	50	42	1013	6	913	96	69	1085
	06:00	26	3689	90	59	3864	15	3793	147	89	4043
	07:00	50	6473	121	49	6692	30	5613	163	68	5874
	08:00	35	4783	190	66	5074	25	4420	252	89	4786
	09:00	28	3938	263	81	4310	23	4204	341	110	4677
	10:00	25	3866	262	73	4226	23	4309	342	100	4774
	11:00	27	3936	251	70	4284	24	4347	331	95	4797
	12:00	31	4138	245	74	4488	28	4584	296	95	5004
	13:00	34	4402	221	74	4731	33	4774	283	95	5186
	14:00	36	4648	216	70	4970	40	5016	277	88	5420
	15:00	70	5681	219	65	6035	67	6160	256	79	6562
	16:00	101	7117	170	65	7453	109	7600	227	79	8015
	17:00	100	6399	141	72	6712	102	6879	192	81	7254
	18:00	50	4475	87	61	4672	45	4824	148	72	5089
	19:00	19	2584	52	49	2705	17	2828	89	61	2995
	20:00	12	1568	36	44	1660	9	1729	55	56	1849
	21:00	6	1196	26	37	1265	5	1290	39	51	1386
	22:00	4	930	23	35	992	5	974	33	45	1058
	23:00	1	573	16	27	618	4	595	25	36	660
Total		**661**	**72111**	**2762**	**1225**	**76759**	**616**	**75712**	**3711**	**1612**	**81650**

4 GANTRY-TO-GANTRY DATA

The gantry-to-gantry data was derived from recording the first entry of a number plate and tracking this number plate through consecutive toll gantries until a specified time period expires without the vehicle passing another gantry, i.e. the vehicle left the freeway. Table 3 contains a small sample of the over ten million records collected during September 2012.

Table 3: Sample of the Gantry-to-Gantry Data From the Source

TRIPID	REGCLASS	STARTDT	ENDDT	STARTTG	ENDTG	DISTANCE	TRAVTIME	AVESPEED	TGCOUNT
137379429	2	2012/08/01 00:00:33	2012/08/01 00:00:33	1019	1019	0	0	0	1
136917450	2	2012/08/01 00:00:34	2012/08/01 00:00:34	1008	1008	0	0	0	1
137342501	2	2012/08/01 00:00:47	2012/08/01 00:10:47	1009	1005	20.1	600	120.6	3
137756442	2	2012/08/01 00:01:57	2012/08/01 00:01:57	1022	1022	0	0	0	1
137845175	2	2012/08/01 00:02:17	2012/08/01 00:09:04	1022	1045	13.3	407	117.641	2
137285672	4	2012/08/01 00:03:24	2012/08/01 00:08:40	1042	1039	8.3	316	94.557	2
136849500	2	2012/08/01 00:03:52	2012/08/01 00:03:52	1001	1001	0	0	0	1
137839755	2	2012/08/01 00:04:31	2012/08/01 00:04:31	1001	1001	0	0	0	1
136984627	2	2012/08/01 00:04:40	2012/08/01 00:10:37	1017	1015	8.6	357	86.723	2
137664113	2	2012/08/01 00:04:46	2012/08/01 00:14:46	1012	1016	18.6	600	111.6	3
136898392	4	2012/08/01 00:05:49	2012/08/01 00:11:35	1029	1031	6.7	346	69.711	2
137872878	4	2012/08/01 00:06:15	2012/08/01 00:12:43	1004	1006	8.6	388	79.794	2
137038339	4	2012/08/01 00:06:23	2012/08/01 00:17:08	1002	1040	15.5	645	86.512	2
137625732	2	2012/08/01 00:06:37	2012/08/01 00:06:37	1002	1002	0	0	0	1
137176338	4	2012/08/01 00:06:42	2012/08/01 00:06:42	1032	1032	0	0	0	1
137655587	4	2012/08/01 00:07:20	2012/08/01 00:17:02	1006	1008	11.3	582	69.897	2
137206507	3	2012/08/01 00:08:36	2012/08/01 00:23:26	1006	1019	18.2	890	73.618	3
137078253	2	2012/08/01 00:08:39	2012/08/01 00:43:59	1002	1012	48.2	2120	81.849	6
137544631	2	2012/08/01 00:09:47	2012/08/01 00:16:53	1032	1025	14.2	426	120	3

Source: ETC Central Operations Centre
Note: Distance measured in kilometres, time in seconds and speed in kilometres per hour.

Tabulation queries and filters enabled the retrieval of a variety of gantry to gantry information, according to:

- Day of the week
- Hour of the day
- Vehicle classification

Where a vehicle only passes through one gantry, the distance, travel time and speed are zero as these measurements cannot be determined from a single point entry.

4.1 Average Speeds

The gantry-to-gantry data provides the average speed for each vehicle travelling between gantries. This data was categorised by time of day and day of the week to provide speed profiles between gantries and average route speeds for the volume-delay functions used to represent the freeways. Figure 2 depicts speed variations by time of day between consecutive toll gantries, each point representing a gantry and the average speed to the next gantry. This graph highlights the variability in the speeds during peak periods along some freeway sections.

The calculated speeds were averaged over a ten kilometre section of freeway and within this distance there are additional on- and off-ramps so the traffic volumes are variable, it is not possible to derive a direct speed-flow relationship. However in traffic modelling this information is valuable for the validation of the volume-delay functions used to represent the freeways as the validation of these functions must be done along routes and not at single points.

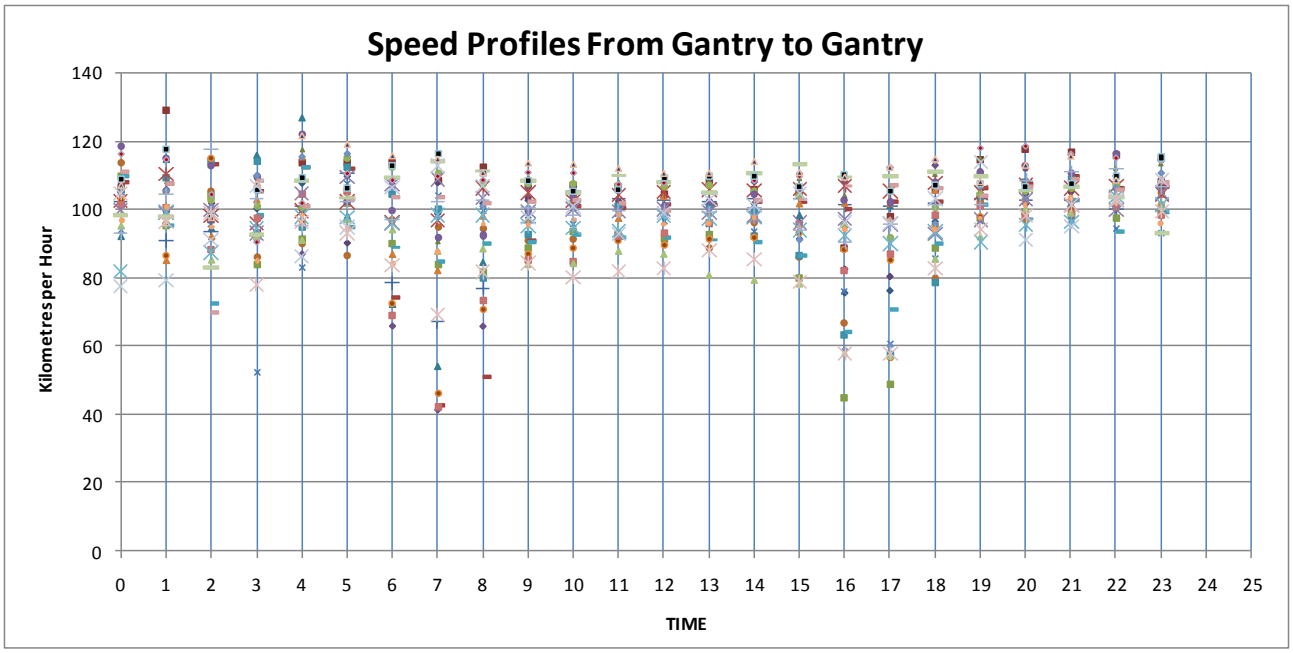

Figure 2: Gantry to Gantry Speeds by Time of Day (Average Weekday)

4.2 Origin Destination Information

The data in Table 3 includes the first and last gantry that a vehicle passed under within a specified time, i.e. indicating an approximation of the trip made by each vehicle on the freeway. Manipulation of the gantry-to-gantry data produced matrices of trips that pass under successive gantries in terms of a gantry origin-gantry destination matrix.
As this is comprehensive and continuous data covering the entire freeway network and it provides continuously updated traffic patterns for further analysis. It will be possible to extract matrices to represent:

- Various time periods including:
  - Individual hours (morning and evening peak hours)
  - Weekday, average weekday and weekends
  - Seasonal monthly, and
  - Annual average daily traffic patterns
- Each vehicle class
- Combinations of the above

Table 5 is one such trip matrix derived from the data set. It represents the over one million data entries in the data set analysed. The values in the cells represent the individual movements between the gantries. The values along the diagonal represent trips that only pass through the one gantry, i.e. start gantry = end gantry.

Note that this does not include short distance trips that use the freeways but do not pass under a gantry.

Table 4: Gantry-to-Gantry Trip Matrices

To / From	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1028	1029	1030	1031	1032	1037	1038	1039	1040	1041	1042	1045	Total
1001	21 207																																					21 207
1002		23 350		11 788		5 130		7 699		1 774		989	1	329		682			656		452	2	543		321	12	6	52	3					2 983	9 448		6	66 226
1003	9 970		22 351																																		32 321
1004				9 979		2 737		3 131		700		399		84		147			368		293		241		95	3	2	20	1								4	18 204
1005	2 616		4 513		11 632																																18 761
1006						11 566		13 855	1	2 393		1 461		579	2	923			1 568		1 676	2	1 208		772	18	8	155	5				1			2	12	36 207
1007	4 361		7 096		15 391		16 308																														43 156
1008								24 362		5 973		3 343	3	922	4	1 000			3 707		4 843	1	2 014		890	34	14	250	8			2				2	125	47 497
1009	913		1 713		2 323		2 400		33 509																												40 858
1010										30 976		12 853		1 506		1 554		1				8			3		78		27								2	47 008
1011	613		826		1 178		1 436		15 957		12 236																										32 246
1012												13 350		4 194		6 168		3		1		31			133		1 184		433								3	25 500
1013	334		498		776		867		6 216		7 196		8 673																								24 560
1014	2		1		1		2							8 054		11 670		7		4		484			641		2 451		2 345								67	25 729
1015	188		108		490		394		940		2 158		3 697		8 267																							16 242
1016	1				2		1									11 410		6		9		491			508		1 400		1 379					3			60	15 270
1017	369		319		731		382		967		1 981		3 048		6 051		22 678	37		70		2 175			647		2 076		2 910					12	9		36	310	44 808
1018	279		326		840		940											9 165																				11 550
1019	12										2		3		10				19 114		21 918		4 424		1 582	131		480				13	18			10	1 804	49 521
1020	288		403		962		1 221											13 758		8 394																		25 026
1021	29				1				1		4		10		75						17 534		4 461		1 777	714		900				30	30			17	712	26 295
1022	487		222		865		796											4 384		2 911		23 309										403	158			305	3 009	36 849
1023									6		9		103		552								13 526		8 598	1 295		2 394										26 483
1024	456		168		770		823		6		37		619		1 809			3 619		1 945		5 874		16 123		3 435		2 687				214	95			167	716	39 563
1025																									6 057													6 057
1028	14		11		39		21		103		1 096		2 761		5 524											30 360												39 929
1029	38		11		42		36											134		132		2 209			1 409		6 286		5 223			44	19			38	247	15 868
1030	1		5		4		8		21		317		1 233		2 728											3 500		7 873										15 690
1031	403		142		536		457											1 679		981		6 167			506				10 380			271	107			110	834	22 573
1032	43		35		192		542		4		7		77		748			12 610		6 107			9 093		2 542	2 811		2 469		42 063								79 343
1037				1 020		505		507		148		75		40		35			69		12		26		15			2			6 653			3 020	7 432			19 559
1038	1 128																															7 420						8 548
1039	1 059																															2 317	1 275					4 651
1040																																		2 547	6 777			9 324
1041																																			12 907			12 907
1042	4 584																															13 389	3 995			7 352		29 320
1045	809																															1 815	617			737	9 741	13 719
Total	50 204	23 350	38 748	22 787	36 775	19 938	26 634	49 554	57 731	41 964	25 043	32 470	20 228	15 708	25 770	33 589	22 678	45 403	25 482	20 554	46 728	40 753	35 536	16 123	26 496	42 313	13 505	17 282	22 714	42 063	6 653	25 930	6 327	8 550	36 564	8 776	17 652	1 048 575

5 TRAFFIC MANAGEMENT USES

The GFIP system collects data from the gantries on a continuous basis. This provides two analysis opportunities, namely that of storing the data for time series analysis and that of real-time information for the monitoring of relative changes in the data.

5.1 Time Series Data Analysis

The continuous collection and storage of the data from the toll gantries will enable the analysis of this data at the various sections on the freeways for the following:

- Seasonal profiles in traffic flows for the conversion of short period traffic counts into an annual average daily traffic volume.
- Average daily traffic flow profiles per day of the week
- Average hourly traffic volumes for each hour of the day.

Over time, trends will be established from the above metrics derived from the data. These trends will assist in ongoing forecasting efforts necessary for the continued development of the road network.

Average hourly volume profiles at various locations along the freeways can be established for various days through the year. These profiles would form the basis for real-time traffic management.

5.2 Real-Time Data Analysis

The comparison of average speeds and time of day enables the derivation of speed profiles that define the variation in speeds through the peak and off-peak periods. Having established these profiles, monitoring traffic flows on the freeways on a continual basis provides a means for the detection of sudden reductions in speeds resulting from non-recurrent incidents.

The speed profiles are representative of sections of freeway between gantries. This information could be used to determine the extent of the impact of an incident as well as the reaction time to restore the normal flow.
Further to the above, tracking the impact of an incident could be used towards the automation of messaging displayed on the Variable Message Signs (VMS) located along the freeway.

5.3 Time Series Data Analysis

The trends that are established from the time series data could provide benchmark data for real time traffic variation analysis. Through the continuous comparison of the benchmark data and the data that is continually streamed from the gantries it would be possible to:

- Monitor variations in traffic volumes passing under gantries. An unexpected reduction in flow could indicate a reduction in capacity as a result of an incident.
- Monitor variations in speeds between toll gantries. This information could be used to monitor the extent of the impact of an incident.
- Monitor the variation in the trip patterns in term of the gantry-to-gantry movements. This variation could indicate the number of trips that divert onto the alternative road network in the event of an incident.

Any significant variation in the above can be relayed to the traffic control centre for further action, which may entail:

- Using the CCTV coverage to verify the cause for the change in flow patterns.
- Establishing the extent of the impact.
- Using the variable message signs (VMS) to relay information to drivers that are beyond the extent of the impact.
- Understanding the alternative routes and common decision points where traffic diverts in the event of an incident.
- As a result of the above, developing incident management plans that could support the alternative road network in the event of an incident.

6 TRANSPORT PLANNING AND MODELLING

Key data required for the development, calibration and validation of a traffic model includes:

- Reliable travel time data for the development and calibration of the road network
- Reliable data on the distribution of traffic through the road network
- Reliable traffic counts

The following describes how ANPR data can be used to satisfy the above data needs.

6.1 Speed Flow Relationships

To obtain speed flow relationships for input into developing volume delay functions (VDFs), single point information is needed. Whilst it may be possible to obtain speeds from the toll gantry equipment, this was not information that was obtained from the ANPR information provided.

Although it was not possible to develop VFDs using the data providing the average speed between gantries, calculated journey times between gantries across the network provides excellent data for the validation of the VDFs used in the traffic model to represent the freeway network. The robustness of the VDFs can be validated by comparing the modelled journey times between gantries for each time period being modelled against corresponding measured times from the gantry data.

6.2 Trip Distribution

To establish the distribution of trips in a traffic model two data sets are required, these are:

- Origins and destinations, and
- A trip length frequency distribution

The gantry-to-gantry matrix provides a distribution of trips on the road network, but does not relate these trips to the actual origin or destination of the trip. Therefore the data cannot be used directly to establish an origin-destination matrix. However, it is possible to validate the distribution of a model assignment by comparing the gantry-to-gantry data with a series of select link analyses. The precise details of this analysis are the subject of on-going research.

Combining the trip matrix provided in Table 5 above and the distances between the various gantry combinations provides a partial trip distribution profile. Again this will only apply to trips that use the freeway network. However the data spans the entire network and therefore provides an accurate account of medium to long distance trip making. This trip distribution data is considered accurate and comprehensive but not for the development of a trip distribution function because the distances do not include the first and last portions of trips that are not on the freeway network. However it may be possible to produce equivalent output from the traffic model to compare and validate the model output. Different functions may be validated for any time period and toll classification of vehicle.

6.3 Traffic Counts

The gantry data provides accurate traffic count data per location and toll class as per the count information for the two sample gantries in Table 2 above. The traffic counts do not however require the ANPR equipment and are essentially equivalent to the currently available Comprehensive Traffic Observations (CTO) data.

Although this is the case, it should be noted that loop based traffic counting equipment is able to distinguish between light and heavy vehicles, can determine short, medium and long heavy vehicles and can determine the number of axles per vehicle. They cannot however count vehicles based on their volumetric classification according to the current open road tolling classifications. This data is however available from the toll gantry equipment.

7 CONCLUSIONS AND RECOMMENDATIONS

In the late 1980s and early 1990s origin-destination information along closed corridors was derived from manual number plate surveys. The results of these surveys were notorious for the small proportion of the data that could be matched up and made sense. Time slots were recorded according to time intervals such as 15 minutes which meant that the calculation of speeds was not possible. Furthermore, such surveys were only conducted over one or two days to produce “representative” OD matrices. If the survey cordon was closed one could derive a partial matrix of external to external trips and obtain information of the internal to external and external to internal trip totals. However the internal to internal trips and the distribution of trips to/from external zones internally was unknown.

ANPR data is essentially very accurate number plate survey data that is collected continuously. Whilst the data is available in real-time, the limitations of this data must be understood and carefully considered, some of these limitations include:

- The location of the gantries (ANPR equipment) does not constitute a closed cordon and therefore for modelling purposes cannot be directly related to modelled traffic zones.
- The calculated average speeds are determined over a distance along which traffic volumes can vary significantly, thus making this data unreliable for the determination of volume delay functions.
- The trip length for the trip length frequency distribution constitutes only a portion of the overall trip, i.e. excludes the distance travelled to the first gantry and from the last gantry and not on the freeway network.

The advantages of using this data include the following:

- The development of time series traffic profiles will, over time, provide bench marks against which real-time traffic flows and patterns can be monitored.
- These benchmark profiles will enable:
  - Early incident detection and verification
  - Monitoring the extent on the impact of incidents on the freeway network
  - Monitoring traffic diversion as a result of incidents and the potential impact on the alternative road network.
- In traffic modelling terms, this is significantly accurate and comprehensive, yet underutilised survey data. Currently it is possible to use this data to:
  - Validate volume delay functions along sections of the freeway network
  - Validate trip length frequency distribution functions through comparisons to specified model outputs
  - Validate the trip matrices by producing equivalent gantry-to-gantry matrices from the model and comparing these to the ANPR data
- Although not ANPR data the gantry equipment produces classified traffic count data based the volumetric vehicle classification system for comparison with the traditional axle-based systems of the CTO counts.

Based on the above, one can conclude that the ANPR system produces accurate data and that this data is currently under-utilized in terms of traffic management opportunities and in the development, calibration and validation on traffic models. It is therefore recommended that:

- Discussions be held with the traffic management teams to establish protocols for the storage, manipulation and monitoring of streamed data to assist in traffic management on the freeway system.
- Continue with the analysis of the ANPR data and matching that which can be produced from the data with that which can be produced from the traffic models for the validation of the traffic models that will be relied upon for the future development of our road networks.

References

Cambridge Systematics Inc.; Travel Time Data Collection; White Paper; http://www.camsys.com/pubs/WhitePaper_OD_TTData_Collection.pdf ; January 31 2012

Friedrich M, Jehlicka P and Schlaich J, 2008. Automatic Number Plate Recognition for the Observance of Travel Behaviour, 8th International Conference on Survey Methods in Transport, France.

Michau, G, Nantes A and Chung E, 2013. Towards the Retrieval of Accurate OD Matrices from Bluetooth Data, Lessons Learned from 2 Years of Data. Queensland University of Technology, Brisbane.

Van Vuren, T Carey, C, 2011. Building Practical Origin-Destination (OD/Trip) Matrices from Automatically Collected GPS Data
. http://abstracts.aetransport.org/paper/index/id/3737/confid/17