Traffic Counts for Strategic Transport Model Validation: What Counts?

Peter Hidas and Frank Milthorpe
Transport Data Centre, NSW Transport & Infrastructure

Abstract

This paper discusses what the most appropriate measures of observed traffic counts are for model validation. As the aim of validation is to compare observed counts with model outputs, one must ensure that the counts used in the comparison represent the modelled travel conditions as closely as possible.

An analysis of hourly traffic counts for a calendar year in the greater Sydney region showed that the "observed traffic counts" may be calculated in several ways, leading to significantly different results. The following recommendations are offered. Travel models are calibrated to reproduce a typical "average day" of certain definition: all days of the year, or weekdays, or workdays, or schooldays only. Therefore, the counts used for validation must be calculated from the same set of days that the model is representing. It is more appropriate to use the median value of the counts rather than the mean, as the mean is more sensitive to a few extreme values that may occur over the year.

This study has also demonstrated that there are significant differences between the model outputs and the link counts for any given time period even from the same input data. These differences are due to the static traffic assignment method used in the strategic models.

Introduction

Travel models are expected to replicate observed conditions within reason before being used to produce future-year forecasts. A standard part of the model validation process is to compare modelled traffic volumes with traffic counts measured on the road network. Modelling guidelines describe the methodology to be used for the validation process in fair detail. The guidelines generally concentrate on the statistical tests and the criteria to be used for the comparison of modelled and observed counts (UK DoT 1997, TRB 2007). Some also describe the data collection methods and specify requirements for the accuracy of the observed counts (FHWA 1997, UK DoT 1997, Schiffer and Rossi 2008). However, little or no details are given on *how to construct the appropriate measures* from the available traffic counts to be used in the process.

The Transport Data Centre, NSW Transport & Infrastructure, (TDC) has recently undertaken extensive analysis of the hourly traffic counts available for 550 directional count locations in the greater Sydney region for the 2006 calendar year in order to develop counts for model validation. The process revealed a number of issues in relation to the preparation of the traffic counts to be used for comparison with modelled volumes. It also highlighted the fundamental differences between traffic counts collected at specific locations and modelled traffic volumes reported from strategic transport models.

This paper describes the lessons learned from this analysis and provides some recommendations as to the most appropriate methods to minimise the effects of the unavoidable differences between modelled volumes and observed traffic counts.

Which days?

For a comparison of observed and modelled traffic volumes to be valid, it is important to ensure that the values reflect the same underlying assumptions. Strategic transport models are generally calibrated to replicate some specific "average day" conditions, for example an average weekday or working day. Therefore, the observed counts must also reflect the average of the same set of days.

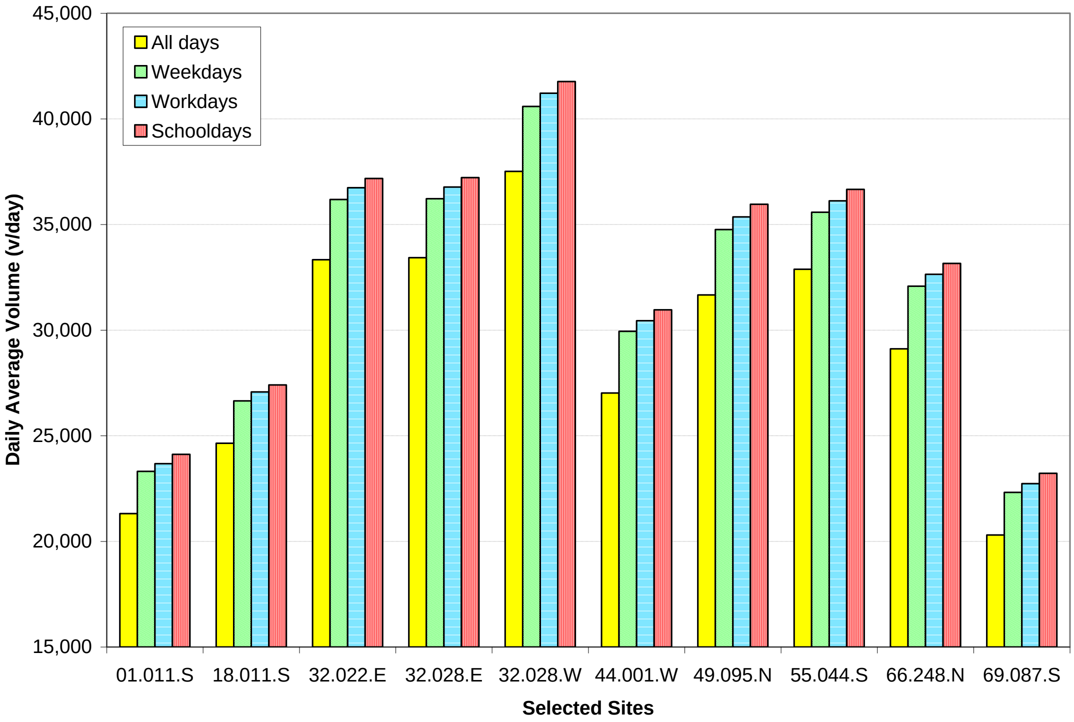

Figure 1 Differences in daily average volumes by day types

Figure 1 shows daily average traffic volumes for a selection of counting stations in the Sydney region, depending on which days of the year are included in the average: all days of the year (365 days), weekdays (260 days), workdays (i.e. excluding weekends and public holidays, 251 days), or schooldays only (202 days). It can be seen from Figure 1 that there are significant differences between the 4 averages: the average of all days is much lower than that for weekdays, workdays is again higher, and the schooldays average is the highest. Averaging over all count locations, and taking the all-days average as reference value (100 %), the following differences were found:

- Average of weekdays: 105.4 %
- Average of workdays: 106.7 %
- Average of schooldays: 107.9 %

If traffic counts are available for the whole year, it is therefore appropriate to calculate the "observed volumes" only from the days that are represented by the transport model. It is also worth noting that even if traffic count data are collected over the whole year in principle, the records may show zero volume for a number of days, due to faults in the counting equipment (loops or tube counters) or in the communication channels. These days should also be excluded from the calculation.

Further complications arise if all or some of the counts are derived from short-term surveys. These short-term surveys are typically conducted over the working days that are also schooldays. In this case, the calculated averages should be further adjusted to account for the differences between the schooldays and workdays, based on available pattern station data, similar to the adjustment used to account for seasonal variation over the year.

Mean or Median?

Once the range of days to be included in the analysis is selected, what statistic should be used to represent the "average day"? While in general we talk about average, i.e. mean, daily (or hourly) volumes, in fact, it is more appropriate to use the median (i.e. the 50th percentile) as a representative of the average volume. Figure 2 illustrates the differences.

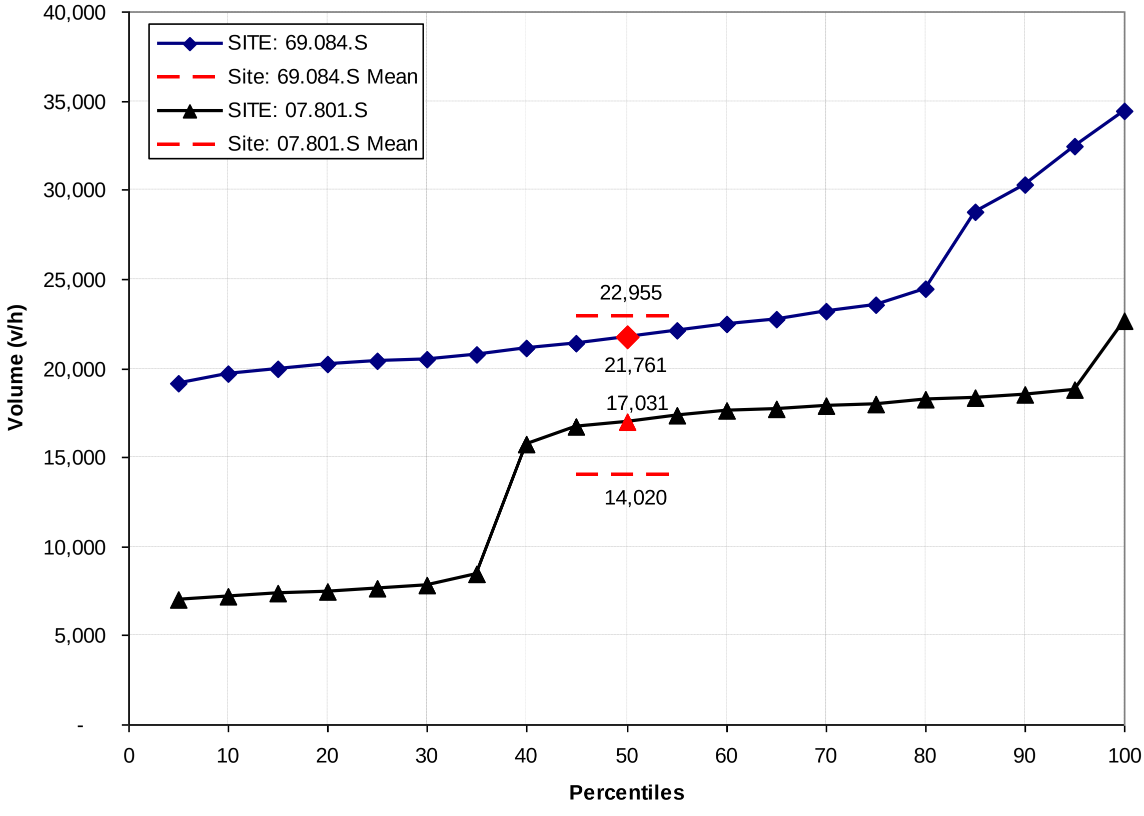

Figure 2 Percentile distribution of daily flows and mean of two selected sites

Figure 2 shows the percentile distribution and the mean of the daily volumes of two selected sites. It can be seen that there are significant differences between the mean and the median. The mean value is more sensitive to a few extreme values that can occur over the year for any reason. Such extreme values can be caused by certain specific events that result in a real surge or decline of traffic activities, for example a cultural or sport festival, a rerouting of traffic from/to elsewhere. On the other hand, they can also indicate an error in the data collection. For example, the sudden change in the percentile distribution at site 07.801.S is most likely caused by a defect in the loop detectors in some lanes of that site. Whether realistic or due to errors, these extreme values are not representative of the conditions that the strategic model is expected to replicate, therefore it is more appropriate to use the median rather than the mean of the observed counts. While the difference between the mean and the median is quite small in most cases, it can vary by 5 percent or more at sites where such extreme conditions occur.

Peak Hours

To validate travel models dealing with various time periods of the day, such as peak and off-peak periods, a suitable average must be calculated for the given time periods from the daily traffic counts. If traffic counts are recorded by hourly (or shorter) intervals, it is a straightforward matter.

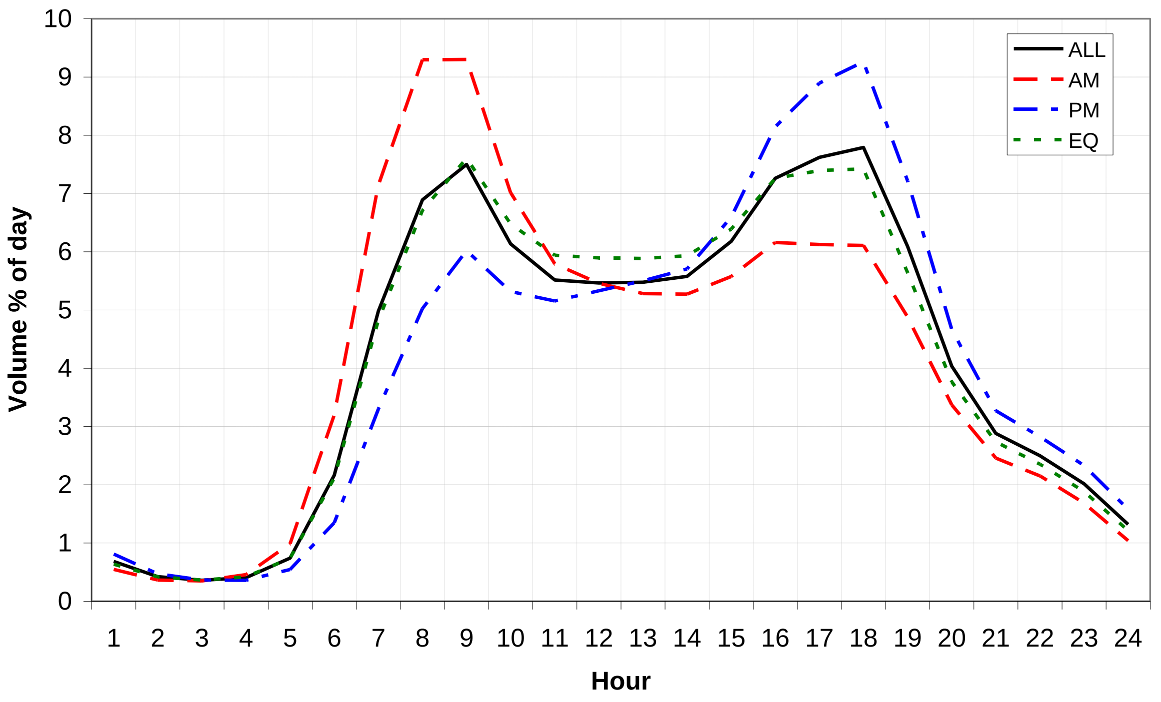

Figure 3 Average hourly profile of daily traffic counts (workdays only)

Figure 3 shows the average hourly distribution of flows as a percentage of the daily total, for 3 groups of sites, depending on the shape of the profile: the group "AM" represents the sites where the morning peak is dominant, "PM" where the afternoon peak is dominant, and "EQ" where the AM- and PM-peaks are relatively balanced. The "ALL" line represents the average of all sites together. From this distribution one can calculate for each site the corresponding values for the time periods used in the travel model. Table 1 shows the average percentages for the Sydney counts.

Table 1 Average distribution of daily traffic counts by time period

Stn. group	AM peak (7 am to 9 am)	Interpeak (9 am to 3 pm)	PM peak (3 pm to 6 pm)	Evening/night (6 pm to 7 am)	Count
AM	18.6 %	34.4 %	18.4 %	28.6 %	207
EQ	14.3 %	36.5 %	22.1 %	27.1 %	85
PM	11.0 %	33.6 %	26.3 %	29.1 %	258
ALL	14.4 %	34.3 %	22.7 %	28.6 %	550

Modelled Trips vs Traffic Counts

It must also be recognised that the number of vehicles observed at each site is not a perfect representative of the number of trips produced by the model for the same sites. Strategic travel models model *trips* and they are calibrated from home interview surveys. In this process, each trip is allocated to a given time period from start to finish, even if some part of the journey occurred outside the period. The allocation can be done by selecting either the departure or arrival time of the trip, or the mid-point of the whole journey. Table 2 shows the differences in the time period allocation factors by the 3 allocation criteria based on TDC's Household Travel Survey, HTS (TDC, 2008).

Table 2 Time period allocation factors from the HTS

Allocated by	AM peak (7 - 9 am)	Interpeak (9 am - 3 pm)	PM peak (3 - 6 pm)	Evening/night (6 pm - 7 am)	Total
Departure time	16.5%	31.5%	26.0%	26.0%	100.0%
Mid-Point time	17.0%	31.3%	25.5%	26.2%	100.0%
Arrival time	16.8%	31.9%	24.5%	26.8%	100.0%

The allocation by arrival time may be beneficial for modelling the morning peak period, but to be consistent over the four time periods of the day, the allocation by mid-point time is the most appropriate. As part of the model validation process, the distribution of counts by time period shown in Table 1 may be compared with the time period allocation factors used in the travel model. There are significant differences between the overall averages in Table 1 and the mid-point time factors in Table 2. However, it is important to note that the number of AM- and PM-dominant sites available in Sydney is significantly different, making the "ALL" average biased towards the PM-dominant group.

Whichever way the trips are allocated to time periods, the assignment results from the model will show as if the trips were present at every count location simultaneously during the time period. This is a consequence of the static traffic assignment method used in strategic travel models. Static assignment does not have a time dimension, it assigns the full demand given for the time period, without considering the time it would take for the vehicles to travel along their selected trajectory. Figure 4 illustrates the differences between traffic counts recorded at specific site locations and assignment results from a strategic travel model.

Figure 4 shows a simplified time-space diagram of travel within the study area. Counting stations are scattered at various locations throughout the area. Trips are represented by their trajectory from origin to destination. In this case we assume that trips are allocated to time periods by the mid-point in time of the journey, represented by the circles at the middle of the trajectory. If the mid-point of the trip falls into the selected time period (8am – 9am in Figure 4), the trip is included in the travel model (shown in the figure by thick lines and red dot at the mid-point), otherwise it is excluded (thin and yellow). The rectangles along the trajectory represent the points where the trip passes a counting station. The dark shaded rectangles represent trips that are included in the modelled time period, the empty ones show the excluded trips. Figure 4 shows that the counts within the modelled time period may include certain trips that are not modelled, like trip A at S4/S5 and trip E at S1/S2, while some trips that are part of the modelled time period are recorded in the previous or the next time period, like trips B and F at S1 and trips D and F at S4/S5 respectively.

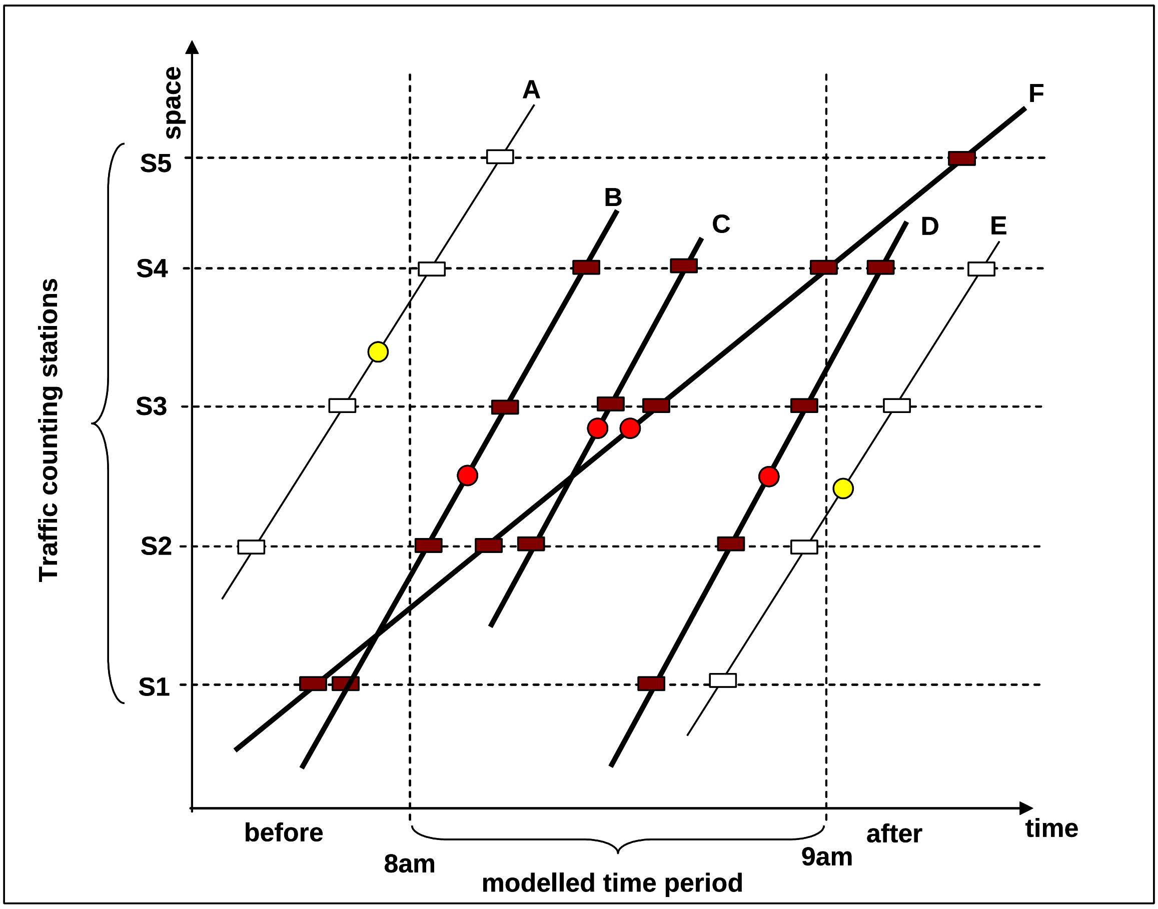

Figure 4 Traffic counts vs modelled trips in a time period

The number of vehicles counted at each station and output by the model is thus different. The counts include all dark and light boxes within the time period, while the model outputs include all dark boxes, regardless of the limit of the time period. As Figure 4 shows, there might be some balancing between counted-but-not-modelled and modelled-but-outside-the-time-period recordings, however, there is no guarantee that the two types of errors would be always equal. The shorter the modelled time period, the more difference can be expected. As far as we know, no information on these differences is available in the literature.

In order to obtain some quantitative measure of these differences, we conducted a detailed analysis of trips using data from the Household Travel Survey (HTS). In the HTS, information is recorded about every trip made during the survey day, including the origin, destination, start time, duration, and mode of travel. From this database we selected all the car driver trips that had any part occurring during a 1-hour period (8 to 9 am). These trips were converted to pseudo-transit lines and loaded onto the Sydney Strategic Travel Model (STM) network in the Emme software platform. As a transit line in Emme, each trip could be easily allocated to an itinerary on the network between the given origin-destination point of the trip. Knowing the start time and the duration of the trips, we could calculate the approximate time at which the vehicle would cross each road section (link) along the journey, allowing us to represent each trip in the time-space dimensions. From this simulation, it is possible to compare the number of trips that would be reported by the travel model for the given time period (based on the mid-point of the trip time) with the number of trips that would be counted as crossing a screenline in any time-slice. To obtain a more systematic picture of the spatial distribution of this comparison, an array of circular rings centred on the Sydney CBD in 5-km intervals were used as screenlines (see Figure 5). Table 3 shows the comparison by ring for the inbound direction of travel.

Table 3 Comparison of model outputs and link counts by ring (inbound)

Ring	Before 8 am		Within (8-9am)		After 9 am		Total		Difference	
	Mod.	Not Mod.	Mod.	Not Mod.	Mod.	Not Mod.	Mod.	Not Mod.	Mod -Count	Percent
5	1,449	11,185	38,813	8,160	5,703	5,241	45,966	24,586	-1,008	-2.2%
10	1,671	14,179	40,756	11,513	5,225	5,981	47,652	31,672	-4,617	-9.7%
15	3,807	17,156	39,306	11,323	2,773	3,946	45,885	32,426	-4,744	-10.3%
20	7,035	17,129	43,959	13,275	2,676	3,893	53,670	34,296	-3,563	-6.6%
25	2,688	13,953	34,260	7,150	2,897	4,685	39,844	25,789	-1,565	-3.9%
30	1,992	15,093	27,912	4,122	2,395	5,589	32,298	24,804	264	0.8%
35	1,434	11,925	18,630	4,367	2,672	3,990	22,735	20,283	-261	-1.1%
40	1,868	9,666	20,223	2,597	1,056	2,913	23,146	15,176	326	1.4%
45	2,330	9,114	19,032	4,376	1,866	2,236	23,229	15,726	-179	-0.8%
50	400	6,301	25,709	4,194	2,366	2,054	28,475	12,549	-1,428	-5.0%
55	1,752	5,669	9,954	562	-	1,593	11,705	7,823	1,190	10.2%
60	1,454	5,419	11,103	860	-	1,071	12,557	7,350	593	4.7%
65	270	3,728	11,558	1,835	-	1,537	11,828	7,100	-1,565	-13.2%
70	430	2,177	14,054	1,162	-	545	14,484	3,883	-732	-5.1%
75	430	2,529	11,273	2,849	255	1,406	11,958	6,784	-2,164	-18.1%
80	1,387	3,822	12,088	1,354	-	1,217	13,475	6,393	33	0.2%
85	2,097	3,053	6,401	985	-	1,217	8,499	5,255	1,113	13.1%
90	784	2,022	5,735	391	343	577	6,861	2,990	735	10.7%
95	567	2,093	3,799	845	140	1,217	4,506	4,155	-137	-3.0%
100	369	2,026	3,384	391	-	1,435	3,753	3,852	-22	-0.6%
105	185	745	3,732	816	-	1,328	3,917	2,889	-631	-16.1%
110	1,075	1,527	3,084	1,567	279	762	4,437	3,857	-214	-4.8%
115	705	1,886	7,883	1,708	896	762	9,485	4,356	-107	-1.1%
120	515	856	5,416	1,287	445	762	6,376	2,905	-328	-5.1%
125	1,010	3,043	8,042	1,026	400	218	9,452	4,287	384	4.1%
130	1,010	1,949	8,185	481	-	218	9,195	2,648	529	5.8%
135	185	1,916	2,578	481	-	-	2,762	2,397	-297	-10.7%
Total	**38,897**	**170,160**	**436,867**	**89,676**	**32,386**	**56,394**	**508,149**	**316,230**	**-18,393**	**-3.6%**

The first 3 double-columns in Table 3 represent the time periods before, within and after the selected (modelled) time period. Each of these is split into two sub-categories: trips that are modelled (Mod.) within the selected time period based on the mid-point time of the trip (refer to Figure 4), and trips that are not modelled (Not Mod.) but would be counted at the screenlines during that time. The Total column shows the sums of modelled and not modelled trips over the full time. The Difference column shows the absolute and percentage difference between the modelled and counted trips for the selected time period itself. This is calculated as the sum of all modelled trips from the whole period, minus the not modelled trips within the time period. The absolute differences vary from a few trips up to several thousand trips. In percentage, the differences vary between +/- 10%. The differences do not show any spatial correlation with the distance from the CBD. This is an indication of the real difference that can be expected between any strategic model outputs and collected link counts in a given time period.

Table 3 presents an overview of the differences between link counts and model outputs at the screenline level. As the positive-negative differences may be more balanced along screenlines, it may also be useful to consider the differences that might occur at individual count locations. Figure 5 illustrates these differences in the Sydney context.

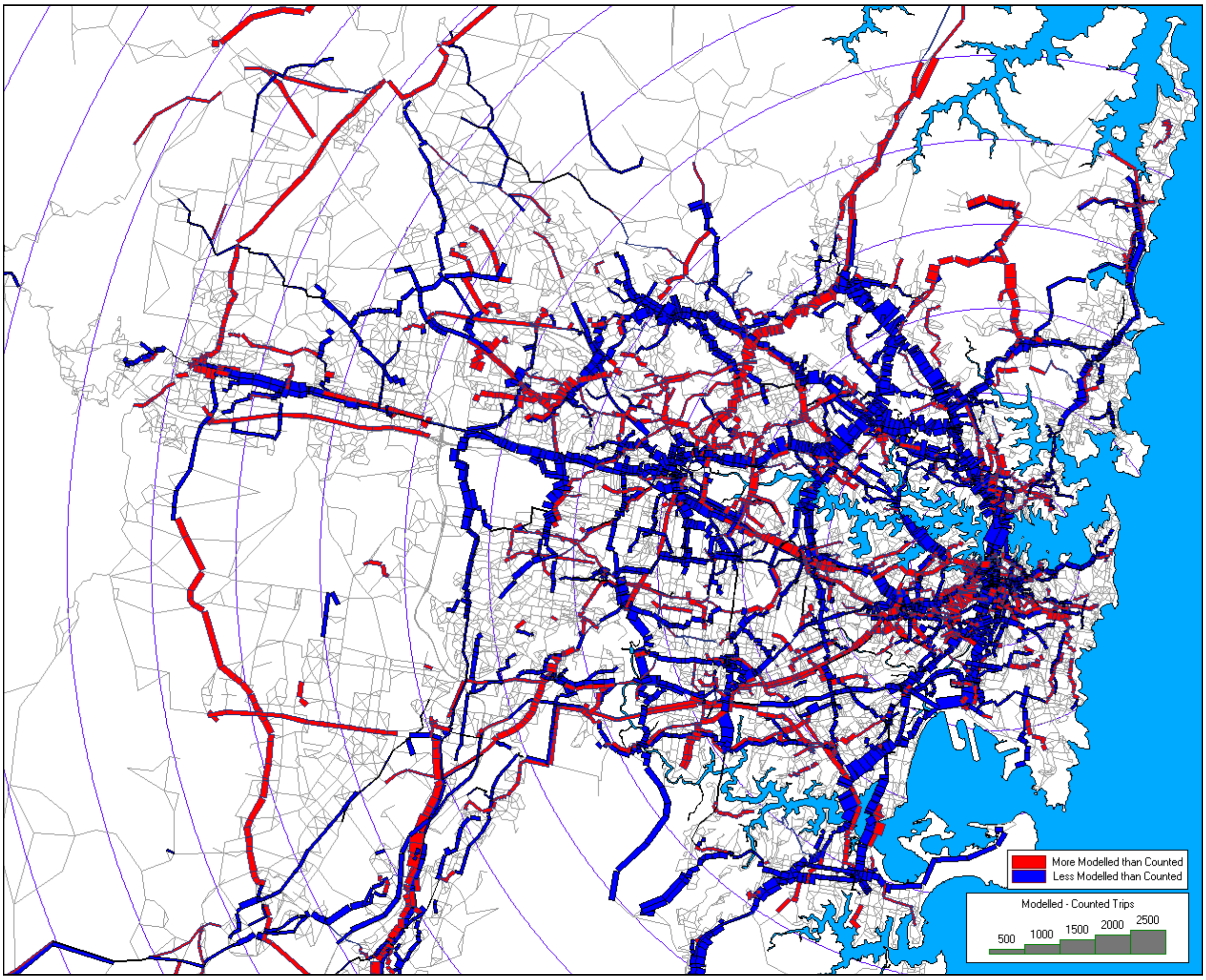

Figure 5 Differences between HTS trips and simulated link counts (8-9 am)

Note that the comparison in Figure 5 is not between actual counts and model outputs. Figure 5 is based on the same HTS trips used in the analyses above, assigned to the network by time categories: before-within-after, and split by modelled – not modelled. The differences shown in the plot are calculated as the sum of (all modelled trips before-within-after) – this is what the model outputs would show for the within time period –, minus the sum of (modelled and not modelled trips within the time period) – this is what a link count would show for the time period. The figure shows significant differences, some positive (red), some negative (blue), often exceeding 1,000 vehicles/hour. It is hard to observe any spatial tendency in the distribution of differences over the metropolitan area. In interpreting these results, one must consider that the HTS trips are based on a small sample survey and factored up to represent the travel behaviour of the whole population, therefore one trip in the survey may represent several hundred trips in reality, using a person weight factor. This means that the trips assigned to the network appear more concentrated on certain links in Figure 5 than they would in reality and therefore the differences shown in Figure 5 are somewhat overestimated. Nevertheless, the figure illustrates that there are significant differences between the link counts and model outputs, due to the nature of the static assignment method used in strategic transport models.

Another important measure of travel activity is the total kilometres travelled (VKT). Table 4 presents the VKT results from the HTS trip analysis, split between the before-within-after time periods and modelled and not modelled trip categories.

Table 4 VKT travelled by time period

	Before	Within (8-9am)	After
Modelled	508,457	7,724,467	598,451
Not modelled	2,128,601	1,293,769	742,980
	Total modelled	8,831,375	
	Total counted within	9,018,235	
	Difference	-186,861	-2.12%

The table shows that a significant portion of the travel modelled within the time period actually occurs in the before and after time period, however, this is more or less balanced by the travel that occurs within the time period but that is not modelled (i.e. the mid-point of the trip is outside the period). Thus the difference in percentage is only about 2 %.

Conclusions

In this paper we concentrated on an important preliminary step of strategic travel model validation: how to prepare the most appropriate measures of observed traffic counts for the validation. This study illustrated that the "observed traffic counts" used for model validation may be calculated in several ways, leading to significantly different results. If the aim of the validation is to compare the observed counts with the model outputs, one must ensure that the counts used in the comparison represent the modelled travel conditions as closely as possible. No information was found in any of the model validation guidelines that we were able to consult on how to achieve that.

Travel models are generally calibrated to reproduce a typical "average day" of certain definition. This may be either an average of (i) all days of the year, or (ii) weekdays only, excluding weekend days, or (iii) workdays only, excluding weekends and public holidays, or (iv) schooldays only, excluding any school holidays. This study has shown that there are significant differences between the average counts of the above groups of days. Therefore, the observed counts used for validation must be calculated from the same set of days that the model is representing. This is an issue not only where counts are available for all days of the year, but also when only short-term counts are available for validation purposes. Short-term counts are typically collected during schooldays, therefore the average of the schooldays must be adjusted to the lower average of the workdays or weekdays that the model represents.

While in general we tend to use the term "average" counts, this study has shown that it is more appropriate to use the median value of the counts rather than the average (mean). The mean is more sensitive to a few high or low values that may occur over the year, either because of some special events or due to errors in the count collection equipment. In either case, the travel model is not designed to represent such unusual events, therefore the observed counts used for validation should be based on the median that is less influenced by these extreme values. This applies to both daily and hourly counts.

If all or some of the counts are derived from short-term surveys, the calculated averages should be further adjusted to account for the differences between the schooldays and workdays, based on available pattern station data, similar to the adjustment used to account for seasonal variation over the year.

This study has also highlighted the fundamental differences between traffic counts observed at specific location at a given point in time and link volumes produced by travel models. Strategic travel models model *trips* and each trip is allocated to a given time period from start to finish, even if some part of the journey occurred outside the period. Travel models allocate trips to the network using a static traffic assignment method. Static assignment does not have a time dimension, it assigns the full demand given for the time period, without considering the time it would take for the vehicles to travel along their selected trajectory. Through a detailed analysis of trips from the Household Travel Survey, this study has shown that there are significant differences between the model outputs and the link counts for any given time period from the same input data. Analysis of different time periods confirmed that the difference is greatest when the time period is short (e.g. 1 hour) and decreasing with longer periods. Unfortunately these differences cannot be eliminated as long as the travel model is using static traffic assignment for the allocation of trips to the network. Nevertheless, it is important to understand the nature and the magnitude of this difference. It represents a practical lower limit of the accuracy that can be reasonably expected from strategic travel models compared with observed traffic counts.

References

FHWA (1997) Model Validation and Reasonableness Checking Manual, Prepared for Travel Model Improvement Program, Federal Highway Administration, Prepared by Barton-Aschman Associates, Inc. and Cambridge Systematics, Inc.

FHWA (2008) Travel Model Validation Practices Peer Exchange White Paper, Prepared for Travel Model Improvement Program, Federal Highway Administration, Prepared by Barton-Aschman Associates, Inc. and Cambridge Systematics, Inc.

Schiffer, R G. and Rossi T. F. (2008) New Calibration and Validation Standards for Travel Demand Modeling, TRB 2009 Annual Meeting (on CD-ROM)

TDC (2008) The Household Travel Survey, Transport Data Centre, NSW Ministry of Transport, http://www.transport.nsw.gov.au/tdc/documents/is2008-03-household-travel-survey.pdf

TRB (2007) Metropolitan Travel Forecasting - Current Practice And Future Direction, Special Report 288, Transportation Research Board Washington, D.C.

UK DoT (1997) Design Manual For Roads And Bridges, Volume 12 Traffic Appraisal Of Roads Schemes Section 1 Traffic Appraisal Manual, UK Department of Transport (http://www.standardsforhighways.co.uk)