DIFFICULT EXPERIENCES AND INNOVATIVE FEATURES OF THE WELLINGTON TRANSPORT STRATEGY MODEL DEVELOPMENT: TRIP DISTRIBUTION, MODE CHOICE, NETWORKS AND MODEL VALIDATION

Authors: David Ashley, Daniel Brown, Andrew Murray, Mandy Liu

1. INTRODUCTION

Our objective in this paper, rather than giving a full description of these elements of the new Wellington Transport Strategy Model (WTSM), is to describe either particular innovative features and some of the practical difficulties which we encountered during model development. We discuss aspects of the development of the distribution/mode choice model in Section 2, the road network in Section 3 and model validation in Section 4.

2. DISTRIBUTION AND MODE CHOICE SUB-MODELS

Both the intended process of estimation and the specification of the distribution and mode choice models underwent great change during the model development process. We discuss why this happened and how we handled the changes. Three aspects are described:

- the statistical estimation software: programs provided with the transport planning packages are poor, but there are some commercially available suites;

- the hierarchical model structure critical to sensible forecasting: this is a deficiency of the present model and has seemingly led to some unsatisfactory forecasting outcomes;

- the treatment of slow modes (that is, walk and cycle trips), a common problem in strategic models.

2.1 ESTIMATION DATA

The estimation travel data was based on the observed data from the household survey, rail and car intercept surveys and a school survey. The table below illustrates the overall mode share and average trip length for each purpose. The observed trip costs (and lengths) were established through the assignment of the observed trip matrices to the base year networks.

◘ **Table 1 Observed Trip Data**

Purpose	Total Trips	Average Trip Length (km's)	Total Trips	Mode Share		
				Car	PT	Slow Modes
Home Based Work	257,000	11.3	257,000	73%	16%	11%
Home Based Education	91,000	6.2	91,000	42%	32%	26%
Home Based Shopping	302,000	4.8	302,000	81%	5%	15%
Home Based Other	368,000	6.1	368,000	79%	2%	19%
Non Home Based Other	481,000	4.7	481,000	70%	2%	28%
Employers Business	154,000	7.5	154,000	87%	1%	12%

2.2 ESTIMATION SOFTWARE

A combination of proprietary and custom written software has been used to develop the distribution and mode choice models. For mode choice we have used LIMDEP, providing the ability to test both multinomial and nested logit models, while for distribution we have implemented a maximum likelihood algorithm with a custom written program, with the advice of John Bates.

2.3 HIERARCHICAL STRUCTURE

There are two general model forms in terms of hierarchy:

- pre-distribution mode choice (encompassing a production zone mode choice model combined with a distribution model segmented by mode), and

- a post distribution mode choice model (with the mode choice occurring at the full matrix level and distribution segmented by car availability).

These two options are presented below.

n **Figure 1 Possible Model Structure for Pre-Distribution Mode Choice**
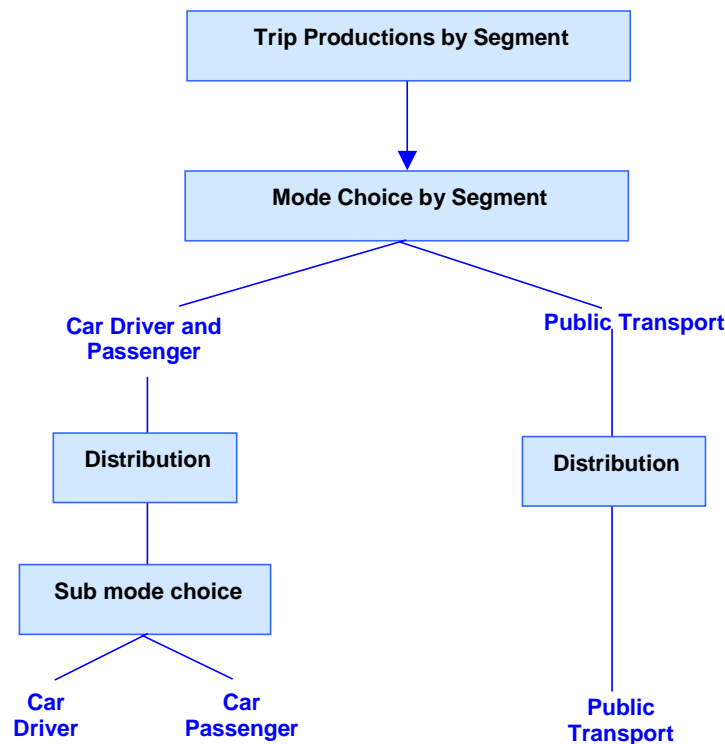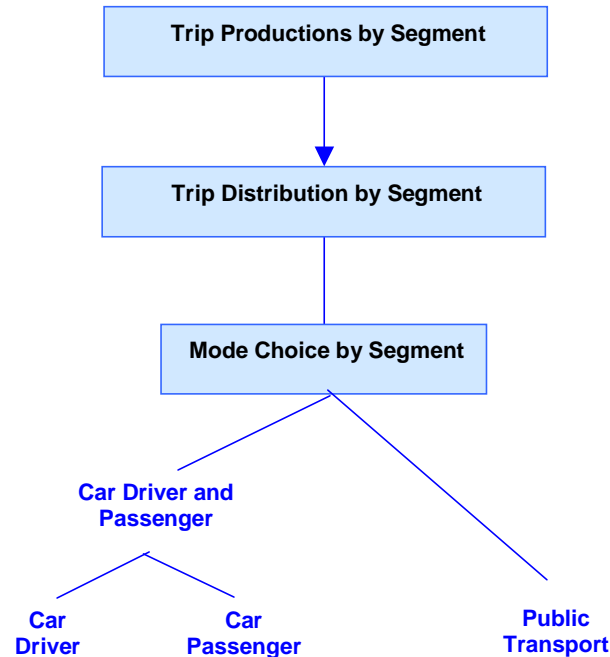

To ensure reliable model sensitivities and forecasts we require model cost parameters to increase as we move down the model structure tree, that is, in the case of the post-distribution mode choice one would expect the mode choice cost parameters to be larger than the distribution cost parameters. In the Wellington context, we investigated both model structures at an early stage to determine the most appropriate form. This process entailed the calibration of preliminary models for

both structures (ignoring slow modes), with a view to adopting that structure supported by the observed data.

It was found that the pre-distribution hierarchies appeared to be best for all purposes except Home Based Work, for which we were unable to calibrate either model hierarchy with any great confidence or consistency. Given the importance of this purpose, this was a setback.

n   **Figure 2 Possible Model Structure for Post-Distribution Mode Choice**

```
          ┌─────────────────────────────────┐
          │   Trip Productions by Segment   │
          └─────────────────────────────────┘
                          │
                          ▼
          ┌─────────────────────────────────┐
          │   Trip Distribution by Segment  │
          └─────────────────────────────────┘
                          │
          ┌─────────────────────────────────┐
          │    Mode Choice by Segment       │
          └─────────────────────────────────┘
                   /              \
          Car Driver and
            Passenger
            /        \                        Public
       Car            Car                    Transport
      Driver        Passenger
```

At this stage we moved onto the final data set and included slow modes in the calibration which introduced a new set of difficulties which we discuss later in the paper. Once these were resolved, we confirmed that pre-distribution mode choice was appropriate for all purposes except HBW, for which neither pre-or post-distribution structures were wholly consistent in the calibrations.

We therefore decided to adopt a simultaneous distribution/mode choice model specification.

In turn this necessitated, the further development of the calibration software. The resulting calibrations were successful and a simultaneous structure was confirmed for HBW. Such a structure is commonly found in international applications, notably of disaggregate models.

2.4   SLOW MODES

The treatment of slow modes (walk and cycle) in the model hierarchy generally poses a number of problems in strategic model calibration, principally because the coarse zone system does not enable the characteristics of walk trips to be measured, most being wholly intrazonal. Our initial plan was to use a structure that drew upon the apparently successful approach adopted by the London Transportation Studies (LTS,

MVA et al 1998). LTS placed the slow mode as a sub-mode of public transport, using distance as a measure of the disutility of walk. This structure was derived after tests combining both slow and public transport and slow and car, with the public transport option preferred.

This method did not work well on the Wellington data, apparently causing the distribution or mode choice model calibrations to fail to converge or to give unsatisfactory results. Investigations strongly suggested that in combining slow with public transport, the resulting composite costs which were fed into the calibration of the distribution and mode choice models had been seriously distorted and were undermining the calibrations of the higher level models.

Recognising the inaccuracy of the walk disutilities and the impacts they were having on the costs being fed to the distribution/mode choice model calibrations, we decided that it was inappropriate to allow the inaccurate slow mode costs to effect the higher level models. It was therefore decided that the costs which are fed into the higher level calibrations should be the unmodified car and public transport network costs. The effect of this decision was to very much improve the calibrations of the higher level models.

We do however retain the representation of slow modes at the lowest level of the model hierarchy. The slow mode trips are combined with either car or public transport for distribution and mode choice, whichever seems most appropriate (discussed below), and extracted at the lowest level in the model hierarchy using fixed proportions as a function of trip distance (ie very short trips are mainly walk).

For each segment (purpose and car availability) the mode with which the data suggested the slow modes were most strongly interacting was selected as the sub-mode with which to combine the slow trips.

An example is provided below of one such analysis for Home Based Work choice. In this example, it is apparent that the car mode share drops most sharply at short distances with slow mode competition. Hence we have combined the car and slow trips.

n   **Figure 3 Home Based Work Choice – Mode Share by Trip Distance**
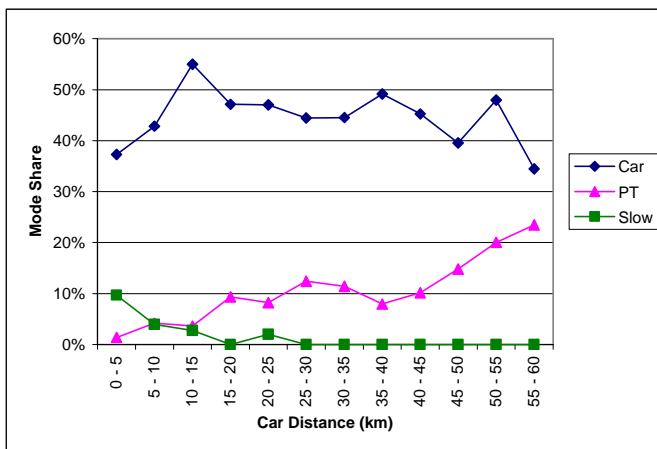
Conversely, the example provided in Figure 4 for Home Based Education combined choice trips illustrates that it is public transport which appears most affected at short distances by competition from slow modes. In this case the public transport trips were combined with slow trips for the trip distribution.

n   **Figure 4 Home Based Education Combined Choice – Mode Share by Trip Distance**
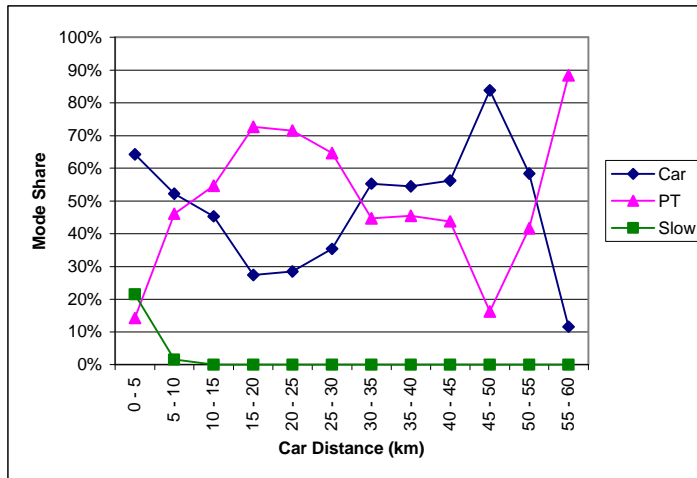
Finally, these figures also illustrate that this is not a perfect judgement but a simplification justified in a situation where slow mode trips are not of paramount importance and cannot be modelled in any detail.

3.   TRANSPORT NETWORK INNOVATIONS

In developing the transport network representation for the Wellington model we have introduced a number of innovations which may be of wider interest to strategic modellers.

3.1   HIGHWAY INTERSECTION MODELLING

In the highway network we have developed a simplified approach to modelling intersections which approximates more detailed project models and at least enables a more sensible representation of the key flow interactions at important junctions across the region to be reflected in the performance of the road network. We have also assembled some evidence that the simple application of speed/flow relationships for a long peak period may significantly underestimate the average travel times experienced by drivers through the peak.

3.1.1   Vehicle Travel Time Functions

The previous version of the WTSM (Booz Allen & Hamilton, 2000) utilised a small number of volume-delay functions to represent all traffic elements of the network, with the different functions representing different capacities and free-speeds. The link capacities were allocated solely on the road type and as such each section of similar road type is treated as a composite link, made up of various traffic control elements (i.e. links and intersections). This implies that all links of the same link type have the same type, size and distribution of traffic control elements (eg. they assume the

same free speed, capacity and size and frequency of intersections). Key limitations of this approach were identified as follows:

- A relatively small number of link types (8) are used to represent a wide range of traffic control elements, resulting in loss of accuracy in travel time predictions where 'non-typical' conditions apply;

- The use of average, composite links parameters makes the predictions at individual control elements within the link inaccurate. These inaccuracies will affect route assignment and travel costs. It also precludes the ability to predict changes to individual elements within the corridor.

The difficulties with the existing approach were highlighted when known or planned upgrades to the network could not be realistically represented and resulted in counter-intuitive delay predictions. An important issue for the update of the WTSM was to address the identified limitations of the current approach to delay modelling, for which the following objectives were defined:

- Provide a strategic level of precision in the modelling of travel delays, consistent with the level of network and zone detail;

- Provide a more accurate representation of key individual traffic control elements (especially key intersections);

- Avoid complex coding systems and the need for extensive network detail; and

- Provide flexibility to use site-specific intersection capacities.

The approach adopted to address these objectives is outlined below:

- Separate the modelling of link travel and intersection delay;

- Use approach rather than turn-level modelling of intersection delay, consistent with the strategic-level accuracy of the network and zone system being used;

- Use a single link and single intersection delay function, but allocate site-specific delay parameters to each link;

- Allocate link capacity parameters based on the road environment;

- Initially allocate fixed intersection capacity parameters based on simplistic indicators differentiating the type and size of the control facility; and

- Allow the capability for intersection capacity parameters at selected sites to be updated during the assignment process (i.e. so capacities are responsive to changing flow patterns).

The previous WTSM used a conic delay function, as defined by Hienz Speiss (1990). A volume-delay function developed by Akcelik (1991) was adopted for the model update because while having a similar form to the Speiss function, it had the advantage of relating to intersection delay modelling and providing explanatory power of the function parameters. The basic function used was as follows:

$$t = t_0 \left\{ 1 + 0.25 r_f \left[(x-1) + \sqrt{(x-1)^2 + \frac{8 J_A x}{Q t_0 r_f}} \right] \right\}$$

where t= average travel time, in seconds per km;
t_0= minimum (zero-flow) travel time;
J_A = Curve Delay Parameter;
x=q/Q = degree of saturation,
T_f = Analysis Flow Period, taken as 1 hour;
q = demand (arrival) Flow rate;
Q = Link Capacity (veh/hr);
r_f=ratio of flow period T_f, to minimum travel time t_0 (r_f=T_f/t_0)

The same function was used to represent both link travel and junction delays. Link capacities (Q), free-speeds (related to t_0) and side-friction parameters (J_A), were allocated to each link in the model based on the identified road environment. At junctions the approach capacities and minimum travel times were either fixed or calculated based on opposing flows at the junction. Approach capacities were calculated from the total flows arriving at the junction and simplified capacity calculations derived from those used in the SIDRA intersection modelling software.

This meant that capacities and hence delays on an approach to a junction were a function of both the intersection geometry and the arriving flows on the other approaches to the junction.

Where there was fine network detail with high frequency of intersections (such as the cbd), fixed rather than calculated approach capacities were used. This was because the strategic nature of the zone system used in the model was not adequate for predicting accurate turning flows, and hence junction capacities.

3.1.2   Flow Peaking Factors

The Akcelik time-dependant delay functions assume that the flow rates within the period are constant, with no peaking and no initial queuing. However, WTSM has 2-hour assignment models and hence the volume delay functions need to predict the average delay for the 2-hour period.

It is clear from the traffic flow data that the traffic flows are not constant across the 2-hour morning and evening peaks and with the non-linear shape of the delay function, the use of an average 2-hour flow rate will not replicate the average of the delays throughout the period.

This variation, or peaking of demand during the modelled period is often represented in other, more detailed models, either through direct calculation of smaller time segments (e.g. OSCADY, TRIPS, SATURN), or by applying adjustment factors to the average period flow rates (e.g. SIDRA). While such techniques are not often applied to strategic models, there would seem little reason not to apply them if they provided more accurate representation of average delays and could be implemented simply and robustly.

A simplistic method was devised where an adjustment factor was applied to the 2-hour demand flows used in the delay functions, so that the average delay would be predicted.

To develop the adjustment factor, for various flow profiles identified throughout the region the delay was predicted using the Akcelik delay function for each 15-minute segment of the 2-hour period. The average of the segment delays was then compared against the delay predicted using the average period flow rate. As expected the use of average period flows underestimated the average delay for the period. This is demonstrated in Figure 5.

The amount of underestimation depends on the peakiness of flows within the period and on the volume/capacity ratio (i.e. on the point on the delay curve). This meant that the adjustment factor would vary depending on the peakiness of the flow profile and on the volume/capacity ratio of the link. A measure of the flow peaking was developed as follows:

$$Peak\ Factor = \frac{average\ flow\ rate\ for\ periods\ which\ exceed\ the\ average\ period\ flow\ rate}{average\ period\ flow\ rate}$$

Peak factors were found to vary between 1.06 (CBD) and 1.16 (Wellington city, outside the CBD).

n   **Figure 5 Comparison of Segment and Period Delay Predictions**
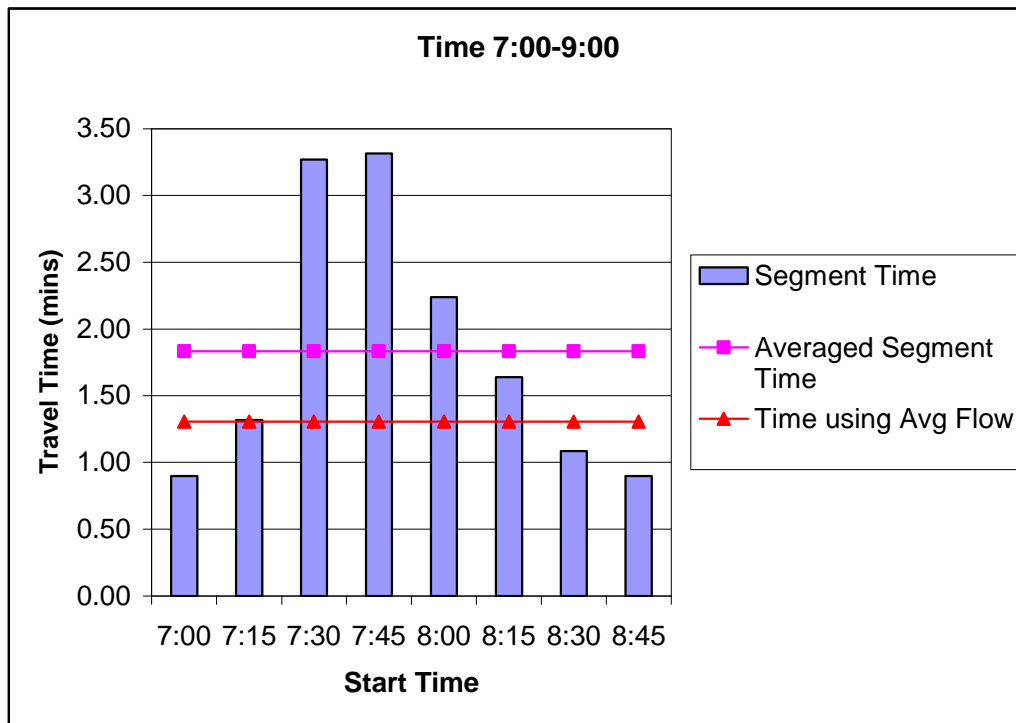
The following figure demonstates the effect of aplying the peaking factor. The curve shown in Figure 6 assumes a free time of 0.75 minutes and a capacity of 3600 veh/2-hours. The volumes on the x-axis are 2-hour flows. Curve 1 includes a peaking factor of 1.1 while curve 2 includes no peaking factor. It can be seen that the use of peaking factor will significantly increase delays, especially at V/C ratios greater than 0.7.

n   **Figure 6 VDF With and Without Peaking Applied**
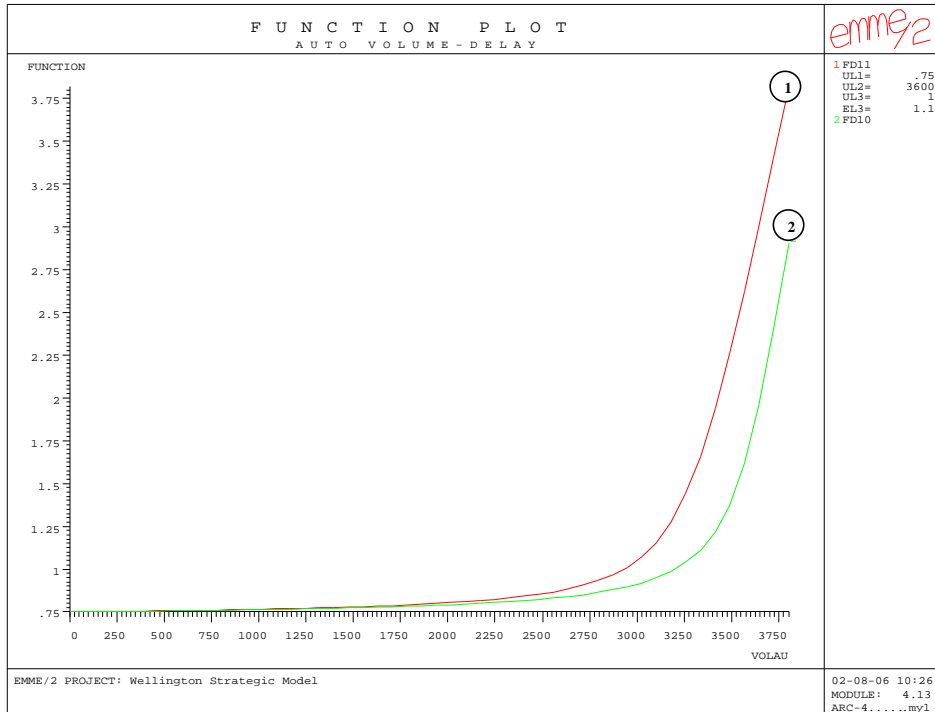
3.1.3   Validation of Travel Time Predictions

The predicted travel times were compared against observed travel times across various routes in the network and a good match achieved. Validation was carried out for all three time periods in the model (am, interpeak and pm), to ensure that the functions also responded appropriately to variations in flows.

3.2   ACCESS TO PUBLIC TRANSPORT

In the public transport network, we paid particular attention to how best to represent access to the rail network, given that this is a key factor in station choice as well as mode choice. The approach we took, described below, is sensitive to the variations in the modes of travel used to access stations and the consequences this has for access times.   While not explicitly modelling park and ride to railway stations the adopted approach does represent the mix of station access modes in determining the station catchment area and access time.

Rail travel data was collected at each railway station in the Wellington Region, encompassing journey origin and destination, access and egress mode, journey purpose and access time from the journey origin to the railway station. This data enabled a detailed analysis of the relationship between access length, speed and mode.

For very short trips, the main access mode used was walk or cycle, however for longer trips the access mode generally switched to the motorised modes, car or public transport. In effect we were able to calibrate a speed function, that reflected

the increase in journey speeds the longer the rail access leg. This function is presented below.

$$Time = (1 - e^{-distance}) * (13.8 + 0.6 * distance)$$

This function was developed by firstly calibrating an access speed function for each mode against distance and then combining these speeds based on the mode share for each distance band.

These rail access links were then added in to the model, connecting each railway station direct to each zone within its catchment area. The links were adjusted so the travel time calculated by the model matched that suggested by our calibrated speed function.

4.   MODEL VALIDATION

One of the most important stages in the validation of models is when the matrices are assigned to the networks and the resulting flows are compared with independent counts. In practice, in complex city models, it is wise to seek a staged process of error identification and removal with only the final stage in this process being the flow validation. In the following we attempt to give a brief insight into the procedures which we followed because they proved to be highly effective in isolating errors and omissions.

A key reason for adopting a highly systematic approach is that errors accumulate through the model system but in calibration no account is taken of this, each sub-model being estimated on survey data rather than using inputs from earlier sub-models in the sequence. What happens therefore is that newly-estimated and apparently well-fitting models show a performance deterioration when they are fed the outputs of earlier models in the sequence. While this is to be expected, the purpose of the structured validation is to identify where losses of performance are significant and to consider how best to correct for them.

The process broadly followed the sequence below, just a few examples of the types of issues which we investigated being given:

- the household surveys and planning data are not necessarily fully consistent, and the incompatibilities would change the forecasts of the trip end models; also the planning data used in the early model calibrations was later updated, and the differences could similarly have affected the trip end model forecasts;

- the estimates of the car ownership model, now slaved to census data, are fed into the trip end model and incompatibities between census and household car ownership data will cause errors in the trip ends;

- errors in the family structure model estimates of the cross-classifications of persons and households will also influence the trip end models;

- the distribution and mode choice models were estimated on observed data and used network costs derived at an early stage in the project:

− synthetic trip ends from the production and attraction models could cause a significant performance deterioration;

− the networks had changed since the original model calibrations and, in addition, the road network speeds would now be estimated in a convergent iterative process; clearly these changes could have a marked effect on the distribution and mode split models;

− at the conclusion of the validation the distribution and mode choice models were indeed re-estimated.

The successful outcome of this process is illustrated by the final model fit figures for the car screenlines (AM, Interpeak and PM periods) and the rail boardings by line (for the AM and Interpeak periods) presented below.
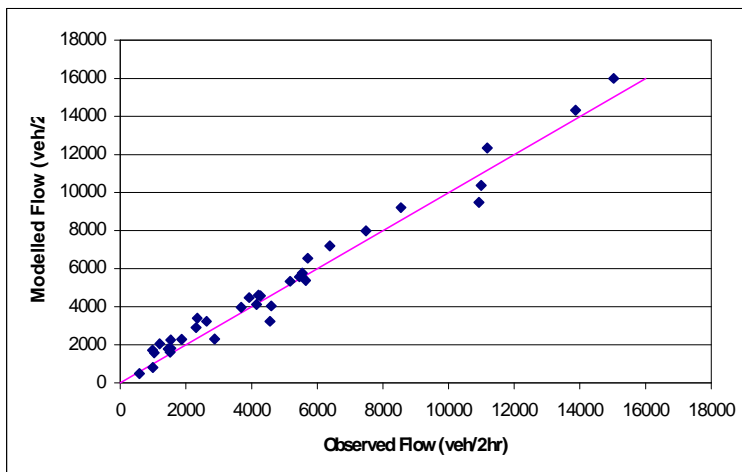
n   **Figure 7 AM Period Screenline Validation**
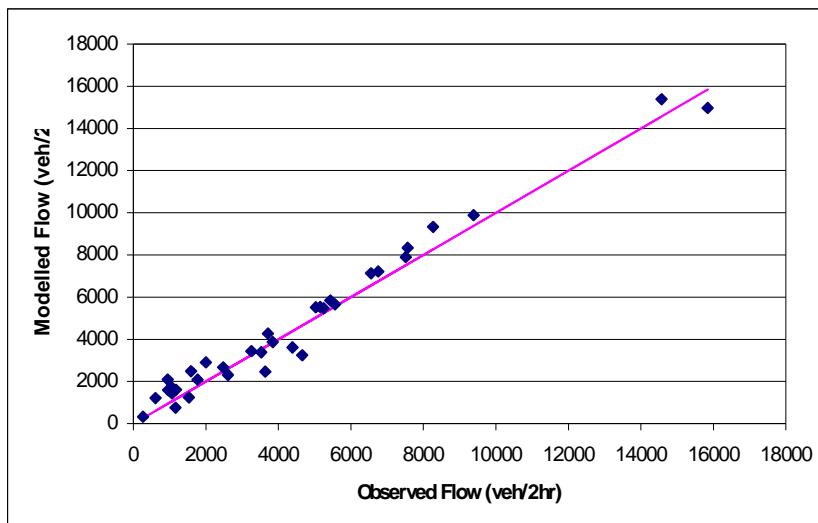
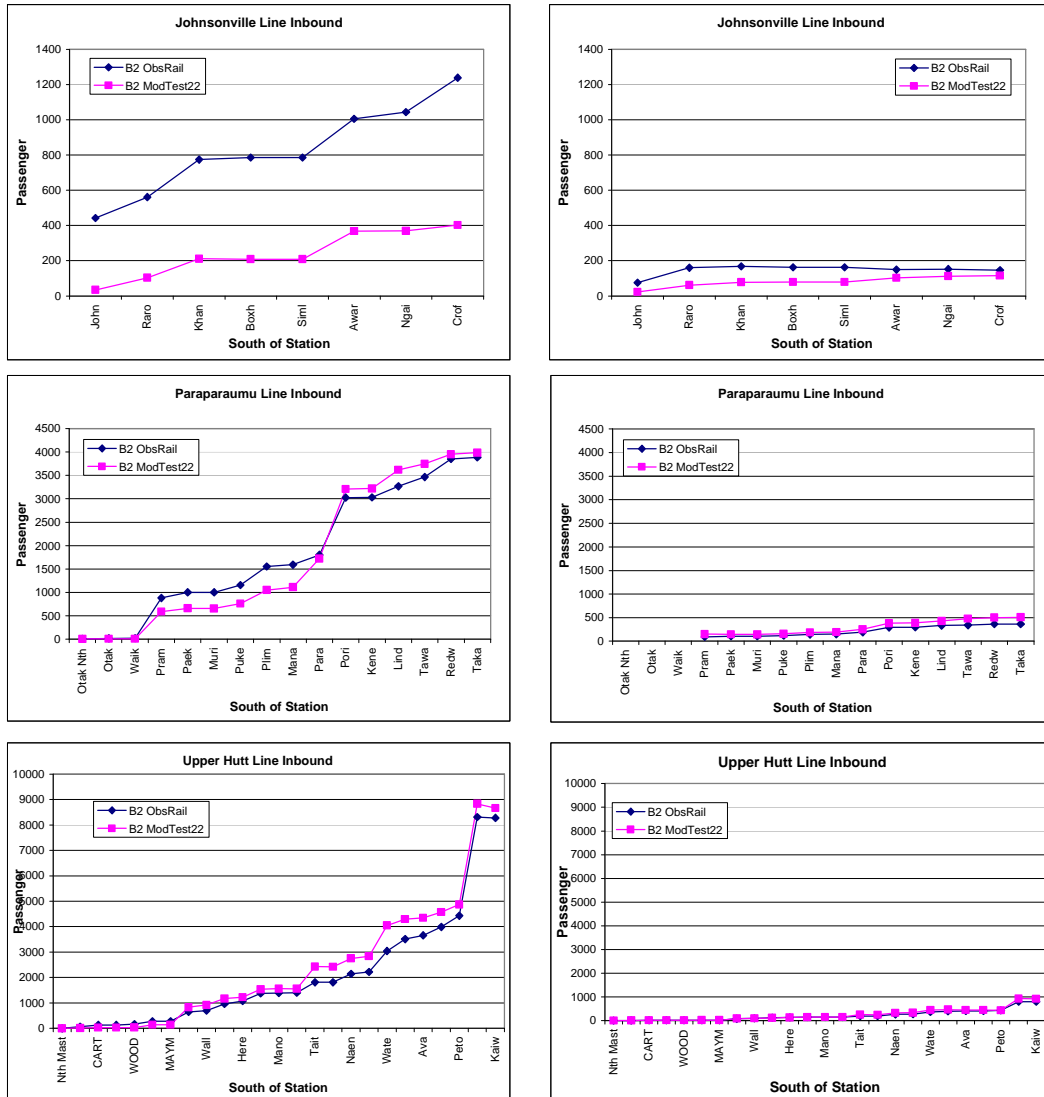n   **Figure 8 Inter Peak Period Screenline Validation**
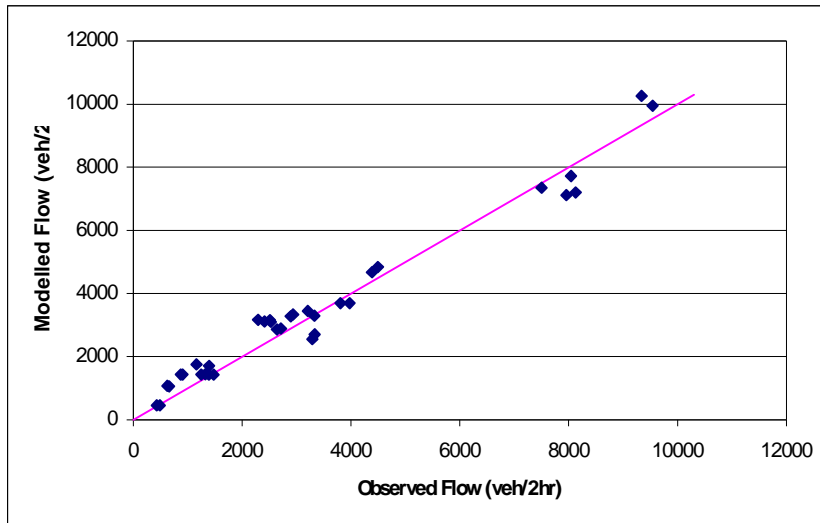
n   **Figure 9 PM Period Screenline Validation**

n **Figure 10 Rail Passenger Loading Validation**

AM Period   **IP Period**

REFERENCES

Booz Allen & Hamilton (2000) *Wellington Area Transport Strategy Model: Draft Report.* Wellington Regional Council.

The MVA Consultancy (1998) *London Transportation Studies - Technical Note 23, Distribution and Mode Split Model Calibration.* Government Office for London.

Speiss, (1990) *Conical Volume Delay Functions.*

Akcelik (1991): *Travel Time Functions for Transport Planning Purposes: Davidson's Functions, its time-dependant form and an Alternative Travel Time Function,* Australian Road Research, 21(3) September 1991.