ESTIMATION OF A PUBLIC TRANSPORT TRIP MATRIX

Transfund New Zealand Research Report No. 149

ESTIMATION OF A PUBLIC TRANSPORT TRIP MATRIX

BOOZ•ALLEN & HAMILTON (NZ) LTD
Wellington, New Zealand

Transfund New Zealand Research Report No. 149

ISBN 0-478-11557-1
ISSN 1174-0574

PO Box 2331, Lambton Quay, Wellington, New Zealand
Telephone (04) 473-0220; Facsimile (04) 499-0733

Booz•Allen & Hamilton (NZ) Ltd. 1999. Estimation of a public transport trip matrix. *Transfund New Zealand Research Report No. 149.* 38pp.

Keywords: bus, matrix, New Zealand, public transport, transport, trips, rail, Wellington

AN IMPORTANT NOTE FOR THE READER

The research detailed in this report was commissioned by Transfund New Zealand.

Transfund New Zealand is a Crown entity established under the Transit New Zealand Act 1989. Its principal objective is to allocate resources to achieve a safe and efficient roading system. Each year, Transfund New Zealand invests a portion of its funds on research that contributes to this objective.

CONTENTS

EXECUTIVE SUMMARY

Introduction

It is difficult to obtain reasonably accurate information on the demand for public transport (PT) in New Zealand cities. Estimates of the PT demand are used to evaluate PT projects, changes to PT services, and changes to PT policy. Better estimates of PT demand will benefit PT users, transport modellers and planners, the road users, tax- and rate-payers, and the New Zealand economy in general.

Objective

The research objective of this project, which was carried out in 1998, was to develop and test a robust methodology for estimating a public transport trip matrix based on the efficient use of a range of existing and accessible public transport information.

Methodology

An algorithm that was derived for this purpose, using the EMME/2 modelling package, is described. The algorithm chosen uses an iterative process to modify a demand matrix obtained from existing data (the 'seed' matrix) until the differences between assigned traffic flows and observed traffic counts on road network links are minimised. At each iteration, a gradient matrix is calculated so that the most significant changes occur on links where discrepancies from the observed data are at their greatest. The process is completed when the objective function (in this case the squared difference of observed and assigned values) has reached a user-specified value.

The algorithm was tested using the Wellington Regional Council (WRC) transport model and its data for the Wellington Region, New Zealand. To test the efficacy of the approach, link data (passenger volumes between stops) were created by assigning a known PT matrix, and then were used in conjunction with a different (but related) initial matrix in an attempt to re-create the known matrix. The method was applied with an observed car driver matrix as the 'seed'. First the car matrix was used without any pre-adjustment and this gave erroneous results. Then the process was modified by scaling the seed matrix to give a correct total number of PT trips, and a much closer estimate was obtained.

The validation was then carried out on a rail service trip matrix, and also for a bus patronage matrix. To complete the validation, the methodology was tested on the combination of the two sets of data (rail + bus).

Conclusions

The most important conclusion is that the performance of the approach is highly dependent upon the quality of the seed matrix.

The basic algorithm does not modify flows unless links are used for which flow data (cordon counts) are available. The importance of ensuring that all significant flows pass through cordons has implications for the choice of locations for the cordons for quarterly transport surveys.

Topics for Further Work

Use of link versus line data

The algorithm adopted could have been applied to counts obtained between each route crossing the cordon individually, rather than to those from the total cordon count. The latter approach however was adopted because limitations in the data, e.g. the incidence of missing or incomplete counts, affected some routes more than others.

By aggregating the data, some information was lost. Further analysis would be desirable to determine the appropriate trade-off in this case.

Use of station boarding and alighting data

Rail data were converted for this project from station boarding and alighting data to link counts. Again this resulted in some loss of information. Further work may be justified to either modify the PT network or to capture this information in another way. EMME/2 calculates boardings and alightings, but does not store this information in a way which would enable comparisons to be made. Otherwise these data could be used instead of the link data. This is a feature of EMME/2 which may not apply to other models.

Use of trip length distribution data to modify seed matrix

The seed matrix is very important to the estimation process. While the literature identifies a general optimisation problem which effectively treats all PT-demand data as equal, the problem has had to be divided into two steps: pre-optimisation and optimisation. This concept could be refined further to incorporate trip length distribution data from ticket sales, to modify the seed matrix.

Solving small PT network estimations

The algorithm uses easily computable data which should enable a small problem to be solved using a spreadsheet model. However a spreadsheet model has not been tested in this project.

Independent data checking

For the matrix adjustment runs, all the link data that were available have been used, and no independent data were available for validation. The comparisons shown, therefore, tend to give an optimistic view of the accuracy of the results. Validation of the data from independent sources would increase the confidence in the performance of the algorithm.

Weighting the data to reflect reliability

Survey data from several sources are often in partial conflict. If this is so, it would be desirable for the method to provide for weighting of the data to reflect its judged reliability. Such weighting is incorporated in the general objective function, but could be included with relative ease.

Analysis of trade-off between data collection costs and accuracy

Increasing the sophistication of the approach is likely to give decreasing returns in terms of accuracy versus cost of data collection. Further work on the sensitivity of the algorithm would enable this trade-off to be established.

ABSTRACT

It is difficult to obtain reasonably accurate information on the demand for public transport (PT) in New Zealand cities. Estimates of the PT demand are used to evaluate PT projects, changes to PT services, and changes to PT policy. Better estimates of PT demand will benefit PT users, transport modellers and planners, the road users, tax- and rate-payers, and the New Zealand economy in general.

The research objective of this project, which was carried out in 1998, was to develop and test a robust methodology for estimating a public transport trip matrix based on the efficient use of a range of existing and accessible public transport information.

An algorithm that was derived for this purpose, using the EMME/2 modelling package, is described. The algorithm was tested using the Wellington Regional Council (WRC) transport model and its data for the Wellington Region, New Zealand, from which conclusions and topics for further work were obtained.

1. INTRODUCTION

1.1 Project Background

This report is of a project concerned with improving public transport demand estimation in New Zealand. Many public transport policy and investment decisions made by transport planners depend on estimating the patronage and revenue implications of the options available.

Estimations of public transport demand are required for evaluating

- Projects (such as Alternative to Roading (ATR) projects),
- Service changes,
- Fare increases,
- Policy changes.

In New Zealand, there is a dearth of good information on true origins and destinations for public transport trips, and this is proving a constraint for transport planning. True origin to destination (O/D) data are important for all the above types of project. However good O/D matrices are expensive to estimate by traditional methods such as surveys.

Initially, the main beneficiaries of the project are seen to be the public sector agencies involved with public transport planning and policy issues. Ultimately, better public transport-demand estimates should feed through to better investment and pricing decisions, and thus they should benefit public transport users, road users, tax- and rate-payers, and the New Zealand economy.

1.2 Project Objectives and Scope

The overall research objective of this project, which was carried out in 1998, was to develop and test a robust methodology for estimating a public transport trip matrix based on the efficient use of a range of existing PT information.

Ideally it will provide a method of providing an initial estimate based on 'expensive' data (such as household interviews), and of updating this data on a regular basis (or as required), using more easily available data such as revenue and passenger trip information.

The project will directly contribute to Transfund New Zealand's objective "*to improve the efficiency and effectiveness of the land transport system by improving demand forecasting*".

The project is seen as having direct applications in:

- Studies of the Transport Disadvantaged,
- Urban and Regional Transport Modelling,
- Alternative to Roading (ATR) projects.

The intention of the project was to develop a methodology that would have a wide application in New Zealand, but at the same time to undertake practical work that could be used and tried by the Wellington Regional Council (WRC).

1.3 Report Structure

The remainder of this report is structured as follows:

Chapter 2	reviews alternative approaches to matrix estimation and outlines the proposed development of the methodology.
Chapter 3	reports on the use of the algorithm with Wellington Regional Council data.
Chapter 4	presents conclusions and topics for further study.

2. DEVELOPING THE METHODOLOGY

2.1 Background

The need for good origin–destination (O/D) data, and the high cost of undertaking the surveys to obtain such data, has led to the development of a variety of techniques for creating matrices[1] for estimating the PT-demand that are based on link data (Van Vliet & Willumsen 1983; Spiess 1987). Link data are much cheaper to obtain, particularly for highway studies where it can be collected by automated counters or manual counts for particular vehicle types.

Most early work on matrix estimation concentrated on solving the O/D data problem for matrices used for highways (Willumsen 1984), although the techniques have since been applied to public transport matrices. Even so, until recently the computation time and data storage requirements of the methods limited their use to very small problems. Increasing computer speed and hierarchical estimation techniques (MVA 1997) now enable very large matrices to be estimated. Passenger trip-matrix estimation approaches have also been developed (e.g. approaches described by MVA 1997, and Spiess 1990[2]), the basic principles being identical to the highway demand case.

2.2 Types of Approach

A review of recent literature carried out for this project indicated that most matrix estimation techniques are variations on the same basic approach. This is to use traffic counts as the primary input, together with a 'seed' matrix, which might be a matrix from a previous survey for example.

The early approaches formulated the problem as an optimisation problem in which the objective was to minimise the difference between the estimated demand matrix and the seed matrix. This approach was subject to the constraint that the assigned[3] volumes from the estimated[3] matrix should correspond to the observed volumes on links with count data.

The method assumed that errors lay entirely in the prior estimate of the matrix, rather than having a component in the traffic counts, and therefore it could be unreliable when errors were from another source.

---

1 A public transport trip matrix ***M*** is defined, with the origin and destination for a trip represented by the rows and columns respectively. Values in each cell $m_{i,j}$ gives the number of trips that are being made from origin zone '*i*' to destination zone '*j*' for a given time period. Typically these matrices are square (i.e. have the same numbers of rows as columns).

2 Information (dated May 1990) on a gradient approach for the O-D matrix adjustment problem by H. Spiess was obtained via an Internet communication.

3 Assigning - technique used to convert data estimated for a trip matrix to link data.
Estimated - data estimated in a trip matrix that represents demand for a PT trip.

2.2.1 General Approaches

More recent approaches have also formulated the problem as an optimisation problem, but used as an objective function the minimisation of the difference between all observed and modelled values. The inclusion of all data in the objective function allows a trade-off between conflicting sources of information (Heydecker et al. 1995).

Following Heydecker et al. (1995), a general form of the objective function is:

$$Z(F, f, w) = \sum_{b \in B} w_b \, S(F_b, f_b) \qquad (1)$$

where:

F	is a vector of model values,
f	is a vector of corresponding data or prior estimates,
w	is a vector of importance weights for the data,
$S(F, f)$	is a numerical measure of similarity between F and f,
B	is an index set of all data used, and
b	is an observation.

The model values F can be expressed in terms of a matrix m_{ij} by the relationship:

$$F_b = \sum_{od} Q^b_{od} \, m_{ij} \qquad (2)$$

where:

Q^b_{od}	is the proportion of traffic from o (origin) to d (destination) present in observation b.

There are three main forms of the measure of similarity $S(F,f)$:

- Entropy,
- Maximum likelihood,
- Least squares.

Each has advantages and disadvantages. Maximum likelihood theory proves that, if this form of objective function can be maximised, then the solution is statistically most consistent with the input information. However it requires a highly sophisticated optimiser, and there is no guarantee that the solution found is a global (rather than a local) optimum.

The maximum entropy and the maximum likelihood approaches minimise the sum of the squares of the relative error, while the least squares approach minimises the sum of the squares of the absolute errors. The least squares approach was chosen for testing in this study because, computationally, it is simpler to implement.

2.2.2 Gradient Approach

All the methods studied solved the optimisation problem using some variation of the gradient approach[2] or method of steepest descent. The method requires a seed matrix, which is assumed to be an estimate, to be adjusted as part of the process.

The gradient approach works by calculating the change in the objective function resulting from a small change in the estimated demand matrix – this is called the gradient matrix – and subtracting a multiple of the gradient matrix from the previous estimated matrix. The multiple is called the 'step length', and is chosen to maximise the reduction in the objective function at each step.

A large number of possible O/D matrices exist that could give rise to a particular observed set of link flows. The gradient method attempts to replicate the observed data with the minimum change to the seed matrix. This is both a strength and a weakness. The algorithm used only adjusts flows for which observed data are available. Also it adjusts only non-zero matrix elements. This has significant implications for the choice of both the seed matrix and the screenline points.

The algorithm works iteratively, with the resultant matrix from each iteration being used as the seed matrix for the next.

2.3 Proposed Method

The method selected for testing in this project defines the objective as being to minimise the sum of the squares of the differences between the observed passenger counts and the assigned passenger counts on selected links. The optimisation problem is solved using a gradient method.

The choice of approach was made on the basis of computational tractability. While more sophisticated methods are available, these require expensive proprietary software. The approach that has been adopted for this project can be implemented with a simple algorithm, which could also be applied manually to small networks.

Because this approach only uses the link data for optimisation, the optimisation function is a special case of the general function (Equation 1), where the weights of all other data are set to zero. However, a feature of the gradient method is that the approach produces a solution which minimises the change from the seed matrix.

As for all gradient methods, the approach requires two types of data:

- Matrix data (i.e. seed matrix),
- Link data (i.e. counts).

The link data are assumed to be the 'correct' data, which the algorithm will attempt to replicate by adjusting the seed matrix. They could be count data collected at selected screenlines or derived from boarding and alighting data.

The problem is thus to minimise $Z(m)$ defined as:

$$Z(m) = \sum (v_a - \bar{v}_a)^2 \qquad (3)$$

where: $\bar{v}_a$ is the observed count on the link a, and v_a is the volume count on the link resulting from the assignment of demand matrix m.

The problem is solved iteratively, with the demand matrix at iteration $n+1$ given by:

$$m_{i,j}^{n+1} = m_{i,j}^{n} \times \left(1 - \lambda^{n} \left[\frac{\partial Z(m)}{\partial m_{i,j}} \right]_{m_{i,j}^{n}} \right) \qquad (4)$$

$$G_{i,j} = \left[\frac{\partial Z(m)}{\partial m_{i,j}} \right] \text{ is the gradient matrix} \qquad (5)$$

and λ^n is the step length for sample n.

The gradient matrix represents the change in the objective function resulting from a small change in the estimated demand matrix. Spiess (1990) has shown that the gradient matrix is the matrix calculated by summing the differences between the observed and the estimated counts along each O/D path, i.e.

$$\frac{\partial Z(m)}{\partial m_{i,j}} = \sum_{k \in K_{i,j}} P_k \sum_{a \in A} \delta_{a,k} (v_a - \bar{v}_a) \qquad (6)$$

where:

- $K_{i,j}$ is the set of valid paths from i to j,
- P_k is the probability of using path k,
- $\delta_{a,k}$ is 1 if a is contained in path k, and 0 otherwise, and

The optimum value of λ (the 'step length') can be shown to be a function of the differences between the observed and assigned trips and the link 'volumes' that result from assigning the gradient matrix *(G)*.

$$\lambda^{*} = \frac{\sum (assign(G) \times (v_a - \bar{v}_a))}{\sum (assign(G))^2} \qquad (7)$$

where:

- λ^* is the optimum step length, and
- *assign (G)* is the 'volume' on link a resulting from the assignment of the gradient matrix G.

Intuitively, because each element of the gradient matrix is the sum of the 'errors' on the links between each origin and destination, the gradient matrix has the desired properties (the O/D paths which cross the screenlines having the largest errors will be adjusted most), but the 'errors' will be counted a multiple number of times.

The optimum step length λ^* can be seen to factor the gradient matrix in proportion to a weighted sum of the errors divided by a weighted sum of the result of assigning the gradient matrix.

2.4 Data Requirements and Availability

The algorithm chosen requires matrix data to provide a starting point, and link data to provide the target observations. It uses an iterative process to determine a suitable matrix by which the difference between assigned traffic flows and observed traffic counts on network links are minimised, based on a demand matrix obtained from existing data. A gradient matrix updates the initial demand matrix through each iteration, by initiating the most significant changes to occur on links where discrepancies from the observed data are at their greatest. The process is completed when the objective function (in this case the squared difference of observed and assigned values) has reached a user-specified value.

2.4.1 Matrix Data

Base information from which an initial matrix could be constructed includes:

- Census data (5-yearly NZ National census),
- Household Interview Surveys (HIS),
- On-vehicle surveys,
- Car driver travel matrices.

Census data provide journey-to-work data obtained at 5-year intervals. It is thus incomplete, is biased to one travel purpose, and is updated only periodically. However in the absence of a full matrix, it could be used as a seed matrix to give approximately the correct 'shape'. It would have to be supplemented by other information to give the right number of trips.

HIS data are often the main source for passenger demand data but, because they are costly to obtain, the sample is usually too small to be reliable for public transport project analysis. Apart from CBD (Central Business District) trips in main centres, public transport typically represents fewer than 5% of all trips, which is less than the usual sample size. Grossing up sampled trips is thus subject to large errors. The proposed technique could be used to make a more robust public transport matrix estimated from HIS data.

Specially commissioned on-vehicle surveys have, in the past, been the only way of obtaining reliable public transport O/D data[4]. Again, these surveys are expensive to undertake, and therefore tend to be undertaken infrequently, or on a sample basis.

The proposed technique could be used in conjunction with on-vehicle surveys in a number of ways:

- The on-vehicle survey could be a sample only, used to establish travel patterns, with cordon count information used to factor up the observations.

---

4 Even these are subject to various sources of error, because such a survey is a set of observations carried out for one day only.

- The on-vehicle survey might only take place on one week day, with cordon counts used to generate matrices for the other days or to estimate a more representative average over the week.
- Infrequent in-vehicle surveys could be supplemented by cordon counts at more frequent intervals, and then the full matrix could be re-estimated.

Car driver matrices are often more readily available than public transport matrices. They may have been estimated from roadside interviews or traffic counts, or from an HIS. Clearly they will be the wrong size (with high numbers of trips compared to PT trips). They will also be the wrong shape because public transport tends to be CBD-orientated, and to have a higher mode share for particular destinations and trip purposes. They may nevertheless provide the best initial estimate of the underlying PT pattern available, and could be used in conjunction with other data (see Section 2.4.3).

The base network could also be estimated from trip origins and destinations and from an impedance function using a standard distribution model calibration approach.

2.4.2 Link Data

The proposed technique requires the observed data to be in the form of link data. Data are not required for every link, for example where the observed flow is entered as 'zero' this is assumed to be 'no observation' and is ignored in the estimation process.

Link data could be available from:

- Cordon counts,
- Station boarding and alighting counts,
- On-vehicle counts,
- Ticket data.

Cordon counts are undertaken regularly by several regional councils in New Zealand. Although in the right format for the proposed algorithm, they may suffer from accuracy problems, as counting passenger numbers is difficult unless the vehicle is stopped for the purpose. If this is done the survey process may impact on the results.

Station boarding and alighting is a feasible way of obtaining counts for rail services. It can also be used at main boarding points to supplement bus cordon data. Also 'ons and offs' are relatively simple to convert to link data. Note however that some of the information is lost in this process. The original 'ons and offs' data can be used to validate the results from the algorithm.

On-vehicle counts are an alternative to cordon counts, and may be undertaken in conjunction with station boarding and alightings.

Depending on the way the system is configured and used, ticket data may possibly be used to extract data on the number of passengers on board at any point from electronic ticket machine (ETM) records.

2.4.3 Other Data

Passenger numbers and revenue (by routes or combinations of routes) may be available from ETM equipment or manual ticket sale data. The operators provide passenger numbers and revenue on a regular basis for each of their contracts. The data are highly aggregated, however, and do not include commercial services (because generally operators are reluctant to make available revenue information which they regard as commercially sensitive).

Even where the data are comprehensive and available to a regional council, they have limitations because they:

- generally give boardings only, and
- relate to journey segments, not to ultimate origins and destinations.

These data are therefore not likely to provide a suitable source for matrix data (although they may provide link data, see Section 2.4.2). They can, however, be used with other matrix data to improve the fit of the seed matrix for use in the algorithm.

Passenger and revenue data could be used to:

- Factor the seed matrix to provide the correct total number of trips (e.g. to allow for patronage growth over time),
- Adjust the trip pattern to match the observed trip length distribution (e.g. where the seed matrix was a car matrix).

2.4.4 Data Preparation

As noted in Section 2.2.2 of this report, a feature of the gradient method as applied is that:

- Flows which do not cross a count point are unchanged,
- Zero value cells in the seed matrix are unchanged.

Depending on the source of the seed matrix, this may or may not be a desirable outcome. In particular, if the seed matrix is an historical matrix that is to be updated, it may be preferable to scale the unchanged cells in the same proportion as those that are modified.

The scaling can be done before applying the algorithm by performing a simple optimisation. This is done to find the scale factor that, when applied to the total matrix, minimises the value of the objective function. The same approach could be used to scale seed matrices from other sources such as journey-to-work or car matrices.

Alternatively, scaling could be undertaken after matrix adjustment by separating the seed matrix into changed and unchanged cells of the matrix, and factoring the unchanged cells by the same average factor as the changed cells.

If the seed matrix originates from a sample survey, the zero elements could possibly be a result of the sampling process. The algorithm can be induced to include these cells by setting their value in the seed matrix to a small positive value.

The method does not provide an efficient means of incorporating all the information available from ticket and revenue data. Generalising the objective function, as is done for the MVEST program (developed by MVA 1997), would compromise the computational simplicity of the algorithm. An alternative approach may be to use trip-length and trip-end information from sales data to adjust the trip length distribution of the seed matrix before adjustment by the algorithm. This is discussed further in Section 4.4 of this report.

3. TESTING THE METHODOLOGY

3.1 Overview

3.1.1 Proposed Methodology

The proposed approach was tested using the Wellington Regional Council's (WRC) transport model. The model is a standard four-step transportation model implemented in EMME/2. The algorithm uses standard EMME/2 procedures for the following functions:

- Public transport assignment,
- Additional options assignment (to sum across O/D paths),
- Matrix calculator,
- Network calculator.

The public transport assignment is multi-path, which should improve the realism of the assignments, and thus improve the performance of the algorithm compared with traditional all-or-nothing assignments. Nevertheless all the above functions would be available in some form with any public transport package, so the use of EMME/2 in no way compromises the generality of the method. Indeed it would be possible to implement the algorithm for a city having a simple radial bus network by using a spreadsheet to hold the matrices, and solving the optimisation problem 'by hand'.

The Wellington Region is divided into 131 zones for modelling purposes. All the analysis was undertaken using 131x 131 element matrices. However in Tables 3.1 to 3.10 of this report, the zones have been aggregated into 10 geographical sectors for presentation purposes.

Testing the algorithm in Wellington Region involved the following tasks:

- Identification of data availability,

- Manipulation of data into an appropriate form,
- Validation tests with 1988 HIS survey data,
- Updating tests to estimated 1996 matrices.

When the algorithm was first used, problems became apparent in replicating observed passenger counts. The problems were a result of discrepancies in the coding of the network. This led to some adjustments in the coding of rail operating speeds in particular. These adjustments have improved the overall performance of the transport model. This 'incidental' outcome is an indication of the power of the technique.

3.1.2 Sources of Data

Possible sources of data that could be used for the seed matrix were:

- The 1988 Household Interview Survey,
- Matrices from Welbus (1993),
- 1996 modelled public transport matrix,
- 1996 national census (journey to work) data,
- The 1988 car driver matrix.

For the validation tests, we started from the car matrix derived from the 1988 HIS, and compared the result with the PT matrix derived from the same source.

For the updating tests, the seed matrix was a combination of the 1988 rail and bus matrices, and estimates were made for the 1996 equivalents.

The results could be tested starting from any of the other sources, but the tests already undertaken have, we believe, been sufficient to draw appropriate conclusions for this study.

Possible data sources for the count information were identified as:

- 1996 screenline data (collected on the 1996 census day),
- 1996 rail boarding and alighting data (collected on 1996 census day),
- Cordon data around Wellington, Hutt and Porirua CBDs (collected quarterly),
- Rail arrivals at Wellington Station (collected quarterly).

In addition, model assignments were used to provide link data for some validation runs.

Other data that could assist were identified as:

- Cars at park and ride stations (regular count),
- Operator returns and ticket sales (obtained from contracted services only).

In the event, only the 1996 screenline, and boarding and alighting, data were used because these were judged to provide the best quality data with which to test the capability of the model.

3.1.3 Data Preparation

The 1996 census data were provided in the form of an EXCEL spreadsheet. This form made it easy for any manipulation that would be needed for model compatibility. Rail passenger count information was supplied on a boarding and alighting basis for each station within the Wellington Region. Bus patronage numbers were given as numbers across defined sites on screenlines. Both the train and bus data were bi-directional.

Because the methodology requires the data to be in a 'link-based' format, both sets (bus + train) of data required conversion. The process was simple but time-consuming because of the amount of data. Typically, however, for continuous updating of the estimated PT matrix, the amount of data to convert would not be as substantial.

For rail, differences between boardings and alightings were calculated to determine the net change in use at each station. The linear nature of the rail network meant that, starting from the outer stations, the net changes could be summed across the rail line until the Central Railway Station. Some of the information contained in the boarding and alighting counts are lost during this conversion. In theory, the estimation methodology could be applied to the on-and-off data rather than the link data. However, there is no efficient way to implement this in EMME/2 version 8.0.

The bus data were not in such a convenient format. Screenline counts were summarised, but location of the counts across screenlines were hard to determine. However, if they could be identified, the information was then possible to transform into the same format as the rail.

Within EMME/2, a link attribute was created to contain the values of the observed rail and bus patronage counts. Where there is no information, and hence the link value is zero, the link is not used in the updating process. Observed zero patronage could be represented by a very small number.

3.1.4 Validation Tests

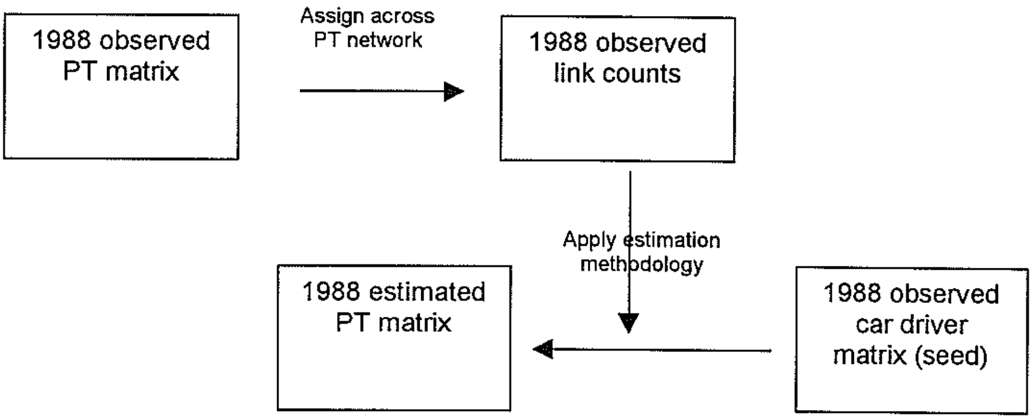
The proposed methodology was assessed in two ways.
First, the algorithm was validated as follows:

- The known 'observed' 1988 trip matrix was taken.
- This was assigned to the 1988 network to give 'cordon counts' at each screenline.
- Using a different matrix as a seed, the algorithm was run.
- The matrix generated by the algorithm was compared with the observed matrix.

Then the updating power of the model was assessed by updating the 1988 public transport matrix using the 1996 rail and bus census day counts. Rail and bus matrices were derived independently and together. There are no 1996 matrix data to verify the matrices against, but the changes in the matrix were checked for consistency with known trends.

3.2 Validation of the Methodology

The validation runs were undertaken using link data from the WRC model created by assigning the 'known' observed matrix, and using a different matrix as the 'seed'. The following is a diagrammatic representation of the process undertaken:



Three different seed matrices were used: car, rail, bus; and then a combination of two sets of data (rail + bus).

3.2.1 1988 Car Seed Matrix

The first matrix tried was the 1988 observed car driver matrix as a seed. Use of the car matrix without any pre-adjustment demonstrated the potential dangers of the approach. The estimated PT matrix 'fitted' the observed data points on the network well, but some cells in the estimated matrix were unchanged by the process. This is a desirable property when the seed matrix has a similar size and shape to the matrix being estimated, but in this case it left values that were clearly erroneous.

The process was then modified by factoring the car driver matrix to give the correct total number of trips. The results of this experiment are shown in Tables 3.1 and 3.2 which are the initial observed car seed and PT matrices, and in Table 3.3 which is the result of applying the algorithm to the scaled matrix.

Comparison of the three matrices indicates that, while the shape of the car driver seed matrix was far removed from what was being estimated, the methodology updates this seed data into an estimate which is not too dissimilar.

Although the seed matrix was modified to be the correct size, the resulting matrix is still too large. A closer result was obtained by further manipulation of the final matrix.

By taking the difference between the seed matrix and the estimated matrix, matrix cells that have changed under the process can be identified. The cells that have changed are the best estimates for the PT matrix. Cells which have not changed are either O/Ds not connected by public transport, or where there is no updating information.

Table 3.1 Observed 1988 car driver matrix (seed matrix used for validation).

Orig. \ Dest.	Eastern Suburbs	Southern Suburbs	Wgtn CBD	J'ville	Tawa	Porirua	Mana	Kapiti	Upper Hutt	Lower Hutt	Total
E.Suburbs	9166	1781	5479	358	56	83	60	51	106	1378	18517
S.Suburbs	1232	3162	3802	281	10	148	0	0	61	729	9426
Wgtn CBD	1190	948	2533	420	96	54	175	166	45	400	6028
J'ville	1223	1174	3128	3682	338	395	35	88	0	1153	11217
Tawa	241	104	740	232	1397	700	96	0	7	417	3935
Porirua	180	313	609	319	350	3349	797	78	53	272	6319
Mana	217	374	1265	163	292	992	1191	261	63	194	5012
Kapiti	81	230	774	60	79	442	445	3024	29	100	5262
Upper Hutt	326	196	670	283	0	47	94	0	7439	2960	12014
Lower Hutt	1208	884	2671	986	153	216	62	0	980	22344	29503
Total	15063	9166	21672	6784	2770	6427	2956	3667	8783	29947	107235

Table 3.2 Observed 1988 public transport matrix.

Orig. \ Dest.	Eastern Suburbs	Southern Suburbs	Wgtn CBD	J'ville	Tawa	Porirua	Mana	Kapiti	Upper Hutt	Lower Hutt	Total
E.Suburbs	2840	491	5671	0	0	0	0	0	32	0	9033
S.Suburbs	606	879	2077	129	0	0	0	0	0	0	3689
Wgtn CBD	888	69	563	82	0	85	130	0	0	239	2056
J'ville	169	443	2732	602	0	0	0	0	0	85	4031
Tawa	0	0	942	0	0	26	0	0	0	0	968
Porirua	80	0	826	11	375	838	152	36	0	0	2319
Mana	0	51	1255	0	60	46	242	157	0	0	1810
Kapiti	0	0	357	0	0	16	0	527	0	0	900
Upper Hutt	0	0	1072	24	0	0	0	0	1399	391	2885
Lower Hutt	0	164	4920	97	0	0	0	0	414	3627	9222
Total	4583	2096	20414	945	435	1011	524	719	1845	4341	36914

Table 3.3 Estimated 1988 public transport matrix using scaled car driver matrix.

Orig. \ Dest.	Eastern Suburbs	Southern Suburbs	Wgtn CBD	J'ville	Tawa	Porirua	Mana	Kapiti	Upper Hutt	Lower Hutt	Total
E.Suburbs	5596	403	5721	23	0	0	0	0	2	10	11755
S.Suburbs	412	2612	2780	131	0	1	0	0	0	46	5982
Wgtn CBD	967	372	1698	12	2	0	0	0	1	11	3064
J'ville	94	975	2421	2652	26	17	0	0	0	127	6311
Tawa	89	112	843	39	973	128	14	0	1	105	2304
Porirua	6	572	133	218	339	1371	201	10	3	18	2871
Mana	3	151	1060	24	130	158	929	173	12	5	2646
Kapiti	0	60	157	92	96	25	125	1700	1	1	2258
Upper Hutt	26	93	762	97	0	0	2	0	4728	587	6296
Lower Hutt	141	2045	2827	147	95	42	1	0	422	8488	14207
Total	7334	7395	18402	3434	1664	1743	1272	1883	5169	9400	57694

These can be updated as follows:

- Compare the observed and estimated matrices, and divide the estimated matrix into two: one consisting of the cell values which have changed (setting all unchanged cells to zero), and the other containing the cell values which have not changed (setting the changed cell values to zero).
- Factor the 'unchanged' matrix so that the sum of the estimated trips matches that of observed trips (other data could be used to further improve this estimate).
- Add the two matrices to obtain the modified matrix.

The resulting matrix is shown as Table 3.4. It has the same general shape as the 1988 observed PT matrix (Table 3.2). Although the absolute values for each O/D are not exact, they do have the same relative behaviour. Accuracy improves with the number of iterations that the algorithm is run.

Table 3.4 Estimated 1988 public transport matrix with modification.

Orig. \ Dest.	Eastern Suburbs	Southern Suburbs	Wgtn CBD	J'ville	Tawa	Porirua	Mana	Kapiti	Upper Hutt	Lower Hutt	Total
E.Suburbs	3333	403	5721	23	0	0	0	0	2	10	9492
S.Suburbs	412	1154	2589	113	0	1	0	0	0	46	4315
Wgtn CBD	967	248	849	12	2	0	0	0	1	11	2092
J'ville	94	975	2421	973	26	17	0	0	0	127	4633
Tawa	89	112	843	16	274	128	14	0	1	105	1583
Porirua	6	572	133	218	339	531	84	10	3	18	1913
Mana	3	151	1060	24	130	120	214	173	12	5	1892
Kapiti	0	60	157	92	96	25	125	448	1	1	1006
Upper Hutt	26	93	762	97	0	0	2	0	1898	587	3465
Lower Hutt	141	2045	2827	147	95	42	1	0	422	4621	10341
Total	5070	5813	17362	1715	964	864	439	631	2339	5533	40732

Figures 3.1 and 3.2 compare the fit of the assigned matrices before and after modification. As can be seen, the algorithm improves the fit significantly. The coefficient of correlation (R2) improves from 0.69 to 1.0, and the least squares line of best fit has a slope of 1 and intercept close to the origin. This extreme test of the methodology results in the observed PT patronage being estimated accurately despite using original numbers that are vastly different.

Figure 3.1 Plot of 'observed' versus 'estimated' matrices before modification.

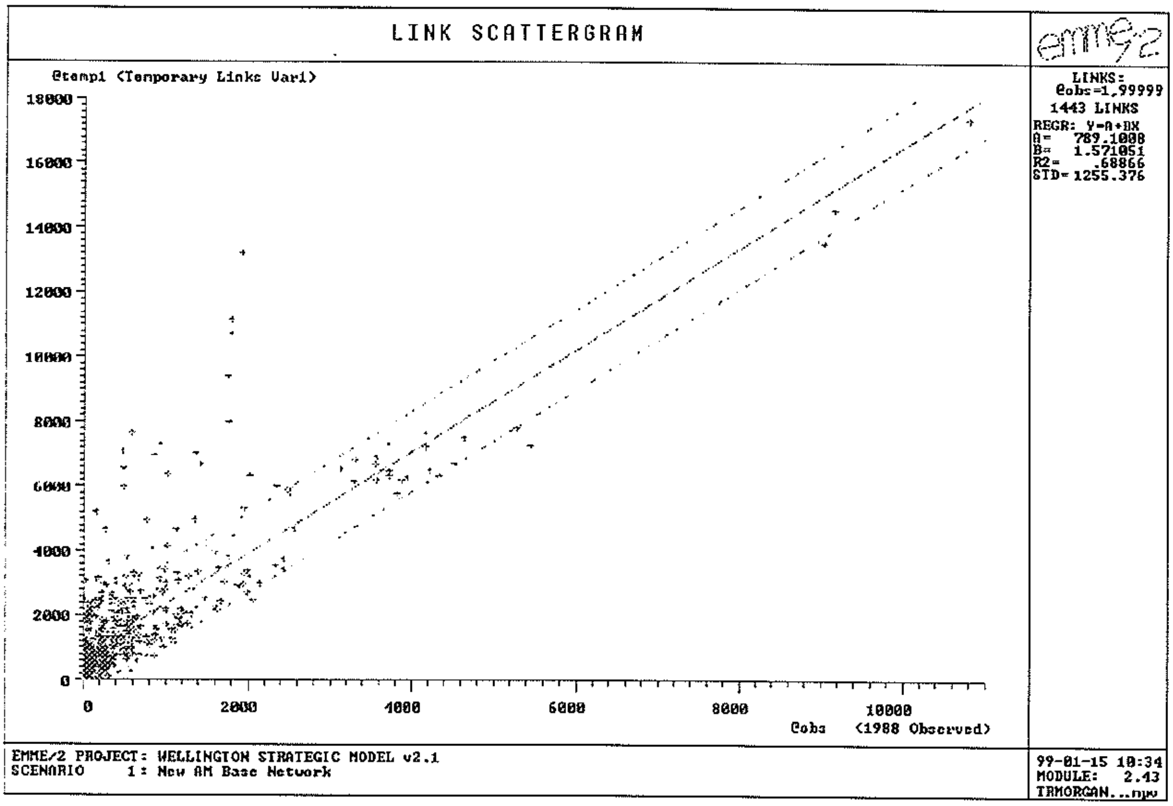
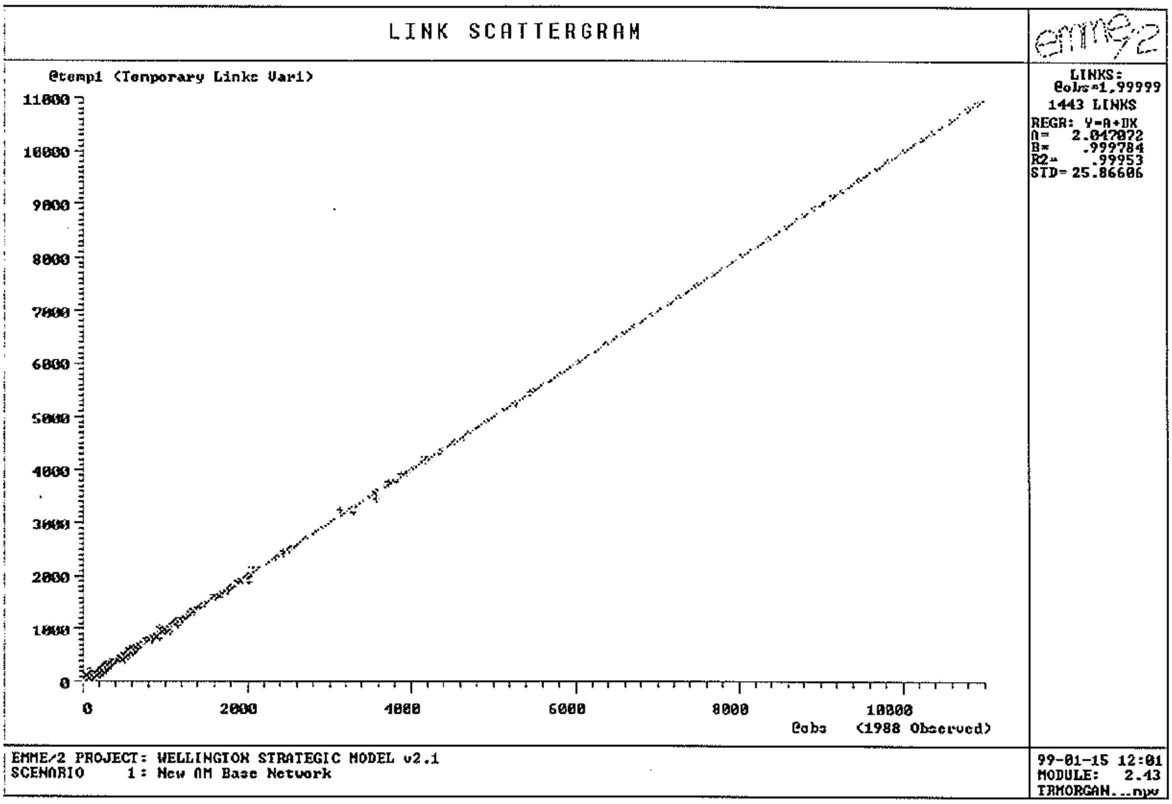


Figure 3.2 Plot of 'observed' versus 'estimated' matrices after modification.



The algorithm was set to run for 300 iterations, and the value of the objective function (the sum of the squares of the errors) after each iteration was plotted. This is shown in Figure 3.3.

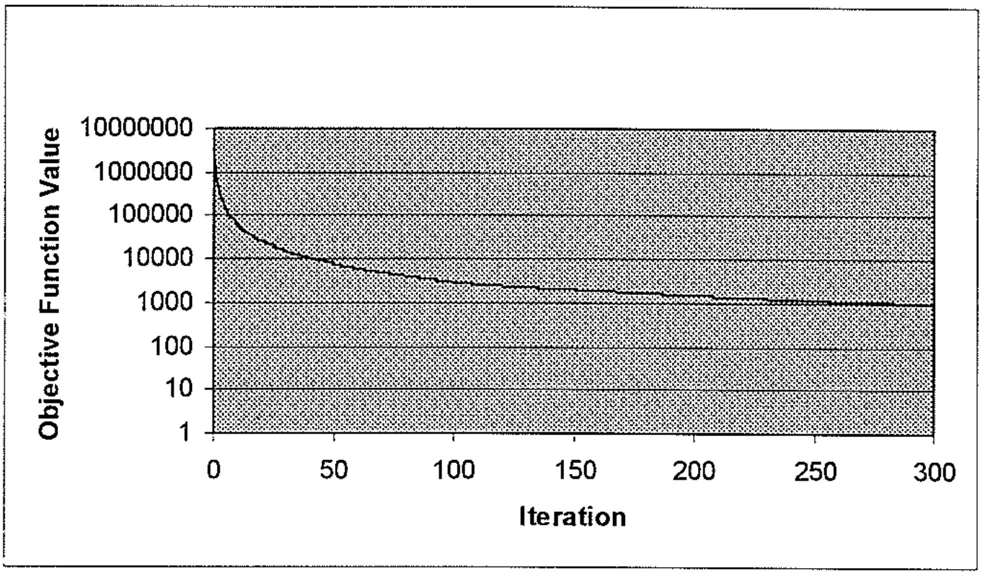


Figure 3.3 Objective function convergence for testing methodology.

As can be seen, the value of the objective function falls rapidly initially and is monotonic (i.e. does not oscillate). These are highly desirable properties. Note that the scale of the y-axis is logarithmic, and that the value of the objective function does not actually reach zero. This results from a combination of interaction between rail and bus during the transit assignment, and anomalies within the specified PT network within the Wellington model.

3.2.2 1988 Rail Seed Matrix

The methodology was applied to the rail services using the 1988 observed trip matrix that was derived from the 1988 HIS, and updating it with count data collected by the WRC on census day in 1996.

Table 3.5 shows the observed morning (AM) peak rail trip matrix generated from the 1988 HIS, while Table 3.6 shows the resulting estimated 1996 trip matrix for the morning peak.

Comparison of the two matrices shows that, as expected, rail patronage decreased over the 1988-96 period. Rail trips from Johnsonville have reduced from 1,902 in 1988 to an estimated 1,181 (38% reduction) in 1996. In contrast, rail trips from the Kapiti Coast has increased by almost 70% over the same period, in keeping with the high population growth that has been observed in that part of Wellington Region.

Rail patronage to the Wellington CBD has decreased significantly by 723, although there appears to be a significant increase in CBD trips from Kapiti. This is in keeping with what was observed by operators. Patronage between Lower and Upper Hutt has been estimated in this study to decrease. During the 1988-96 period, the development of a number of commercial bus services could have had a detrimental effect on rail patronage.

Table 3.5 Observed 1988 morning peak rail trips.

Orig. \ Dest.	Eastern Suburbs	Southern Suburbs	Wgtn CBD	J'ville	Tawa	Porirua	Mana	Kapiti	Upper Hutt	Lower Hutt	Total
E. Suburbs	0	0	0	0	0	0	0	0	0	0	0
S. Suburbs	0	0	197	129	0	0	0	0	0	0	326
Wgtn CBD	0	0	26	82	0	85	130	0	0	239	561
J'ville	0	119	1409	375	0	0	0	0	0	0	1902
Tawa	0	0	942	0	0	26	0	0	0	0	968
Porirua	0	0	826	11	375	24	0	36	0	0	1273
Mana	0	51	1167	0	60	46	31	157	0	0	1510
Kapiti	0	0	324	0	0	16	0	99	0	0	439
Upper Hutt	0	0	1072	24	0	0	0	0	86	358	1540
Lower Hutt	0	0	4357	97	0	0	0	0	369	767	5590
Total	0	169	10319	717	435	196	161	292	456	1364	14109

Table 3.6 Estimated 1996 morning peak rail trips.

Orig. \ Dest.	Eastern Suburbs	Southern Suburbs	Wgtn CBD	J'ville	Tawa	Porirua	Mana	Kapiti	Upper Hutt	Lower Hutt	Total
E.Suburbs	0	0	0	0	0	0	0	0	0	0	0
S.Suburbs	0	0	216	80	0	0	0	0	0	0	296
Wgtn CBD	0	0	26	39	0	62	96	0	0	119	342
J'ville	0	54	854	273	0	0	0	0	0	0	1181
Tawa	0	0	1166	0	0	26	0	0	0	0	1192
Porirua	0	0	744	6	184	14	0	36	0	0	984
Mana	0	35	1010	0	44	37	27	157	0	0	1310
Kapiti	0	0	602	0	0	41	0	99	0	0	741
Upper Hutt	0	0	865	7	0	0	0	0	113	221	1207
Lower Hutt	0	0	4113	44	0	0	0	0	319	586	5062
Total	0	89	9596	449	229	180	123	292	433	926	12316

Figures 3.4 and 3.5 compare the fit of the assigned matrices before and after adjustment. As can be seen, the algorithm improves the fit. The coefficient of correlation (R2) improves from 0.98 to 0.99, and the least squares line of best fit has a slope of 1 and intercept close to the origin.

Figure 3.4 Plot of observed 1996 versus assigned 1988 rail trip counts.

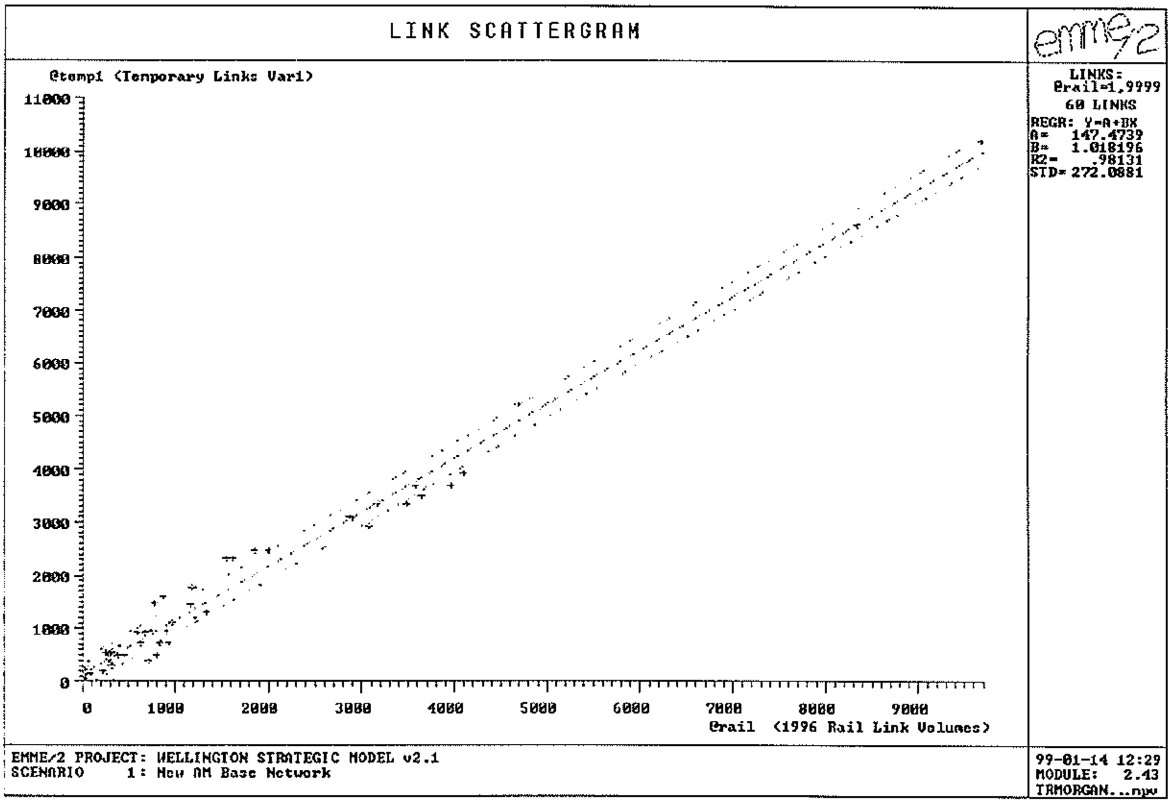
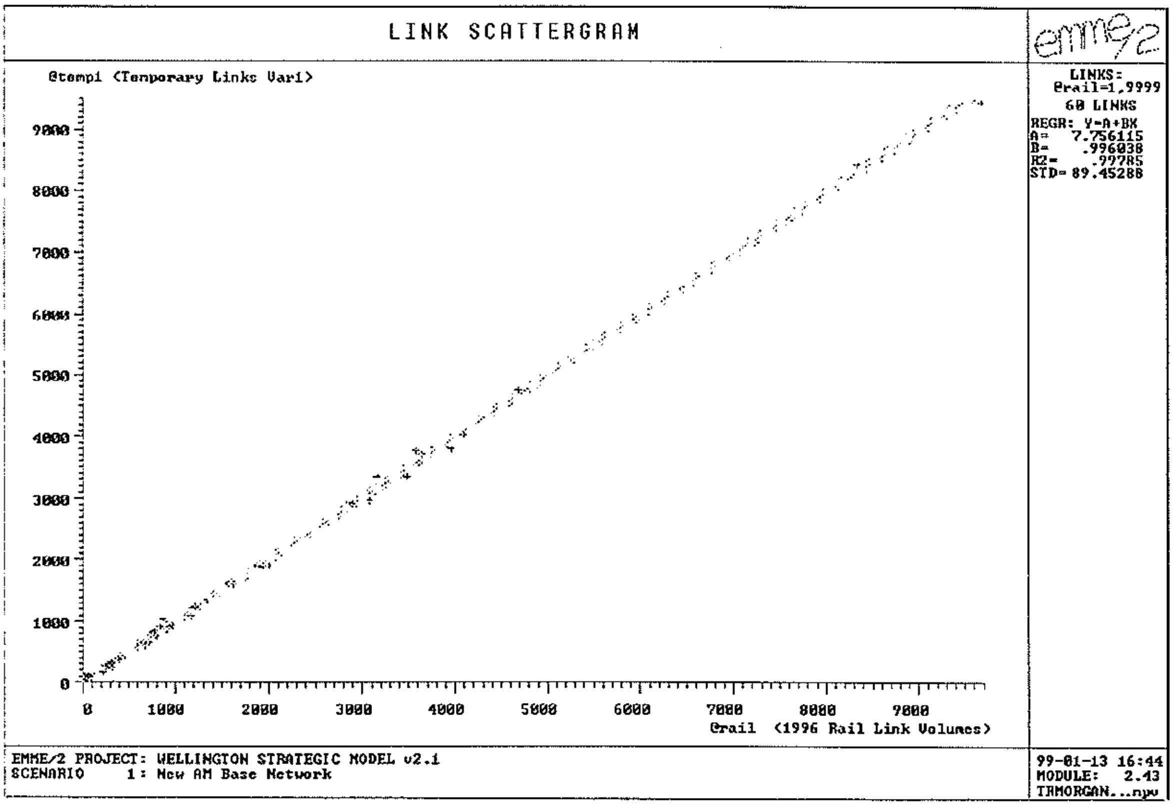
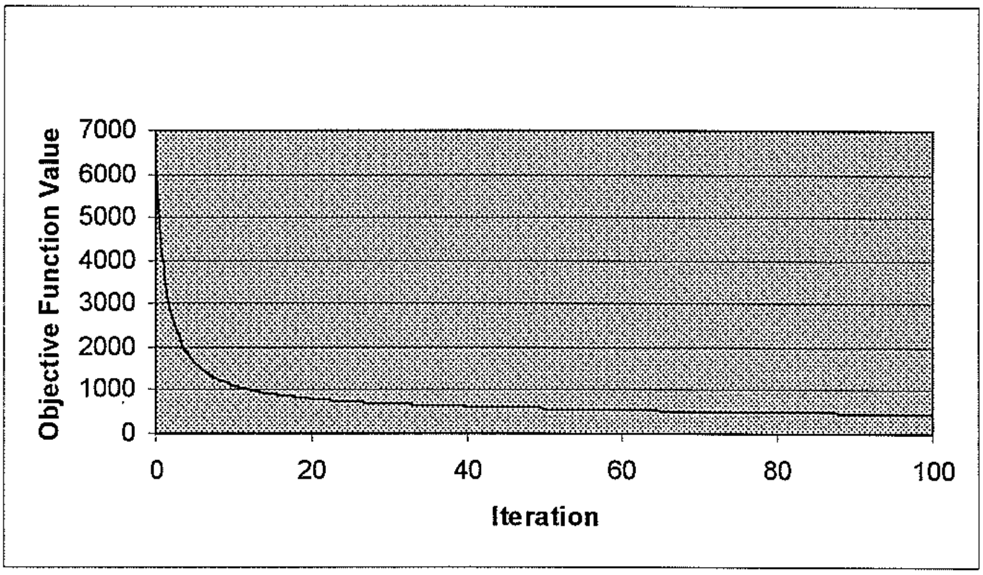


Figure 3.5 Plot of observed 1996 versus assigned 1996 estimated rail trip counts.



The algorithm was set to run for 100 iterations and the value of the objective function (the sum of the squares of the errors) after each iteration was plotted. This plot is shown in Figure 3.6.

Figure 3.6 Objective function convergence for estimation of rail matrix.



As can be seen, the value of the objective function falls rapidly initially and is monotonic (i.e. does not oscillate). These are highly desirable properties. The value does not actually reach zero. This results from the specification of the network. The residual is equivalent to 0.3 persons per cell of the matrix.

3.2.3 1988 Bus Seed Matrix

The methodology was then applied to the observed 1988 bus patronage matrix (Table 3.7). A 1996 patronage matrix was estimated using the 1996 census data (Table 3.8). The total patronage does not appear to have significantly changed over the 1988-96 period. Significant effects are an estimated increase in total bus trips from Lower Hutt (29%), and a decrease in bus trips to the Wellington CBD.

A population increase within Miramar and Seatoun over this period has contributed to an increase in bus patronage within that part of Wellington Region.

Internal patronage within the Lower Hutt area has increased by 11% because of an increase in bus accessibility through the development of new commercial services covering Lower Hutt, Wainuiomata, and Eastbourne during the time frame (1988-1996) studied. Some of these services have been extended to Courtenay Place in Wellington CBD, and therefore their attractiveness of using the bus for trips to the city has increased.

Table 3.7 Observed 1988 morning peak bus trips.

Orig. \ Dest.	Eastern Suburbs	Southern Suburbs	Wgtn CBD	J'ville	Tawa	Porirua	Mana	Kapiti	Upper Hutt	Lower Hutt	Total
E.Suburbs	2840	491	5671	0	0	0	0	0	32	0	9034
S.Suburbs	606	879	1880	0	0	0	0	0	0	0	3364
Wgtn CBD	888	69	538	0	0	0	0	0	0	0	1495
J'ville	169	324	1323	227	0	0	0	0	0	85	2129
Tawa	0	0	0	0	0	0	0	0	0	0	0
Porirua	80	0	0	0	0	814	152	0	0	0	1046
Mana	0	0	88	0	0	0	212	0	0	0	300
Kapiti	0	0	33	0	0	0	0	428	0	0	461
Upper Hutt	0	0	0	0	0	0	0	0	1313	32	1345
Lower Hutt	0	164	564	0	0	0	0	0	45	2860	3632
Total	4583	1927	10096	227	0	814	364	428	1390	2977	22806

Table 3.8 Estimated 1996 morning peak bus trips.

Orig. \ Dest.	Eastern Suburbs	Southern Suburbs	Wgtn CBD	J'ville	Tawa	Porirua	Mana	Kapiti	Upper Hutt	Lower Hutt	Total
E. Suburbs	3700	367	4739	0	0	0	0	0	25	0	8830
S. Suburbs	444	1171	2131	0	0	0	0	0	0	0	3747
Wgtn CBD	648	56	538	0	0	0	0	0	0	0	1241
J'ville	100	198	800	227	0	0	0	0	0	28	1354
Tawa	0	0	0	0	0	0	0	0	0	0	0
Porirua	34	0	0	0	0	814	152	0	0	0	1001
Mana	0	0	88	0	0	0	212	0	0	0	300
Kapiti	0	0	33	0	0	0	0	428	0	0	461
Upper Hutt	0	0	0	0	0	0	0	0	1331	32	1363
Lower Hutt	0	534	940	0	0	0	0	0	45	3185	4704
Total	4927	2326	9269	227	0	814	364	428	1401	3245	23000

Figures 3.7 and 3.8 compare the fit of the assigned matrices before and after adjustment. As can be seen, the algorithm significantly improves the fit. The coefficient of correlation (R2) improves from 0.63 to 0.93, and the least squares line of best-fit more closely resembles a diagonal.

Figure 3.7 Plot of observed 1996 versus assigned 1988 bus trip counts.

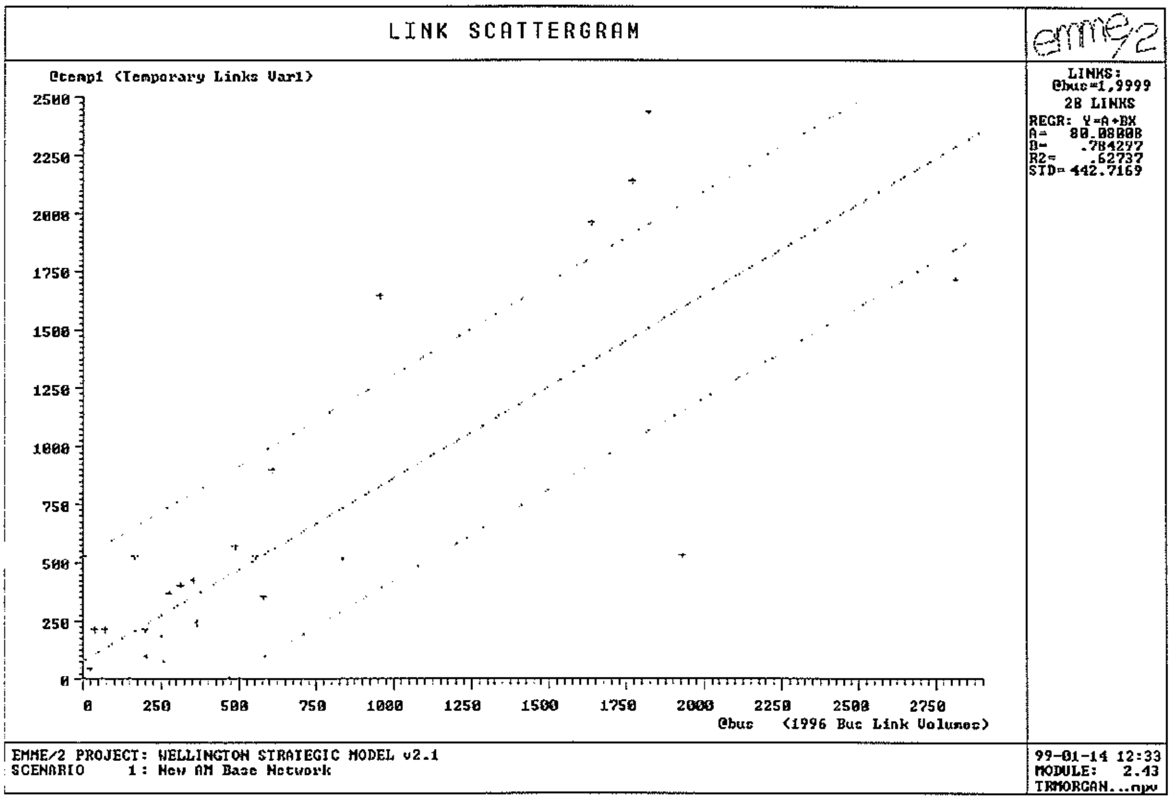
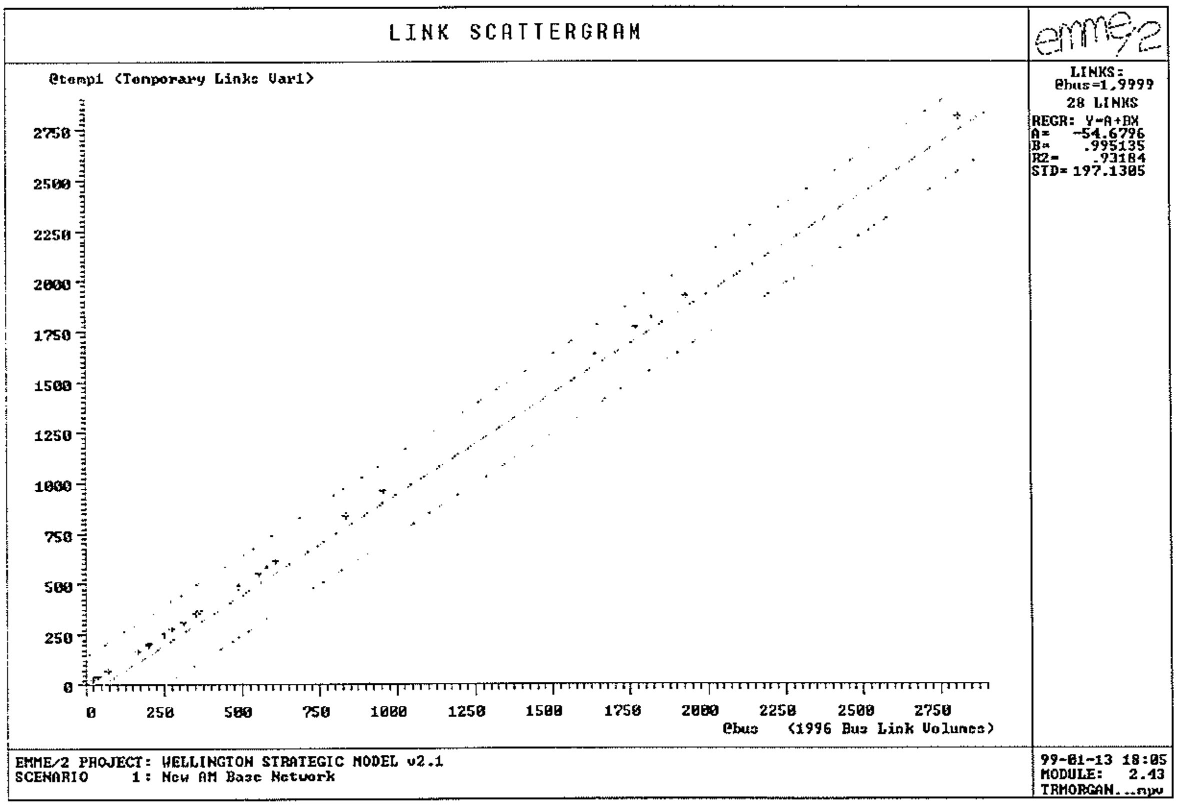
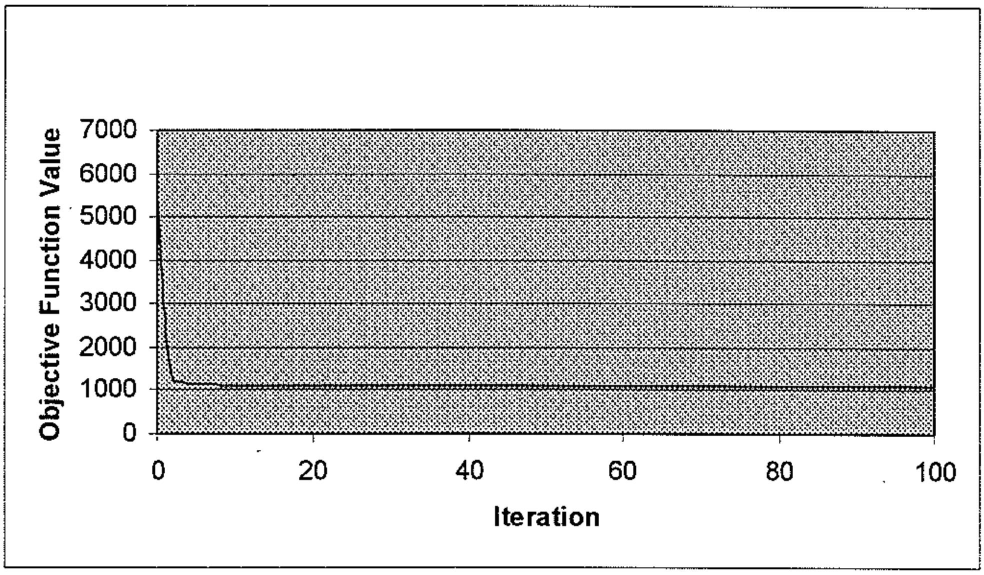


Figure 3.8 Plot of observed 1996 versus assigned 1996 estimated bus trip counts.



As previously, the convergence of the objective function was tested. The estimation was conducted over 100 iterations. Compared with the rail estimation, the bus convergence appears to occur more quickly. This is because of the nature of the bus network, where bus trips can take a choice of a number of routes, compared to the linear nature of the rail network.

Figure 3.9 Objective function convergence for bus trip estimation.



3.2.4 Bi-modal Estimation

The methodology was then tested on the combination of the two sets of data (rail + bus). The seed matrix used was the 1988 observed morning peak total patronage matrix. The total observed matrix (Table 3.9) is the sum of the observed rail and bus matrices. However the estimated matrix (Table 3.10) is not the sum of the rail and bus estimates because of the interaction within the transit assignment of EMME/2.

Overall, total public transport patronage is estimated to have decreased by 7% over the 1988-1996 period. The effects within the matrix are typically the sum of the effects that were estimated by the transit modes independently. The changes are an increase in total transit patronage from Kapiti (especially to the CBD), an increase in trips from Lower Hutt, a large decrease in trips from Johnsonville (especially to the CBD), and general decreases in other parts of Wellington Region.

Table 3.9 Observed 1988 morning peak public transport trips.

Orig. \ Dest.	Eastern Suburbs	Southern Suburbs	Wgtn CBD	J'ville	Tawa	Porirua	Mana	Kapiti	Upper Hutt	Lower Hutt	Total
E.Suburbs	2840	491	5671	0	0	0	0	0	32	0	9033
S.Suburbs	606	879	2077	129	0	0	0	0	0	0	3689
Wgtn CBD	888	69	563	82	0	85	130	0	0	239	2056
J'ville	169	443	2732	602	0	0	0	0	0	85	4031
Tawa	0	0	942	0	0	26	0	0	0	0	968
Porirua	80	0	826	11	375	838	152	36	0	0	2319
Mana	0	51	1255	0	60	46	242	157	0	0	1810
Kapiti	0	0	357	0	0	16	0	527	0	0	900
Upper Hutt	0	0	1072	24	0	0	0	0	1399	391	2885
Lower Hutt	0	164	4920	97	0	0	0	0	414	3627	9222
Total	4583	2096	20414	945	435	1011	524	719	1845	4341	36914

Table 3.10 Estimated 1996 morning peak public transport trips.

Orig. \ Dest.	Eastern Suburbs	Southern Suburbs	Wgtn CBD	J'ville	Tawa	Porirua	Mana	Kapiti	Upper Hutt	Lower Hutt	Total
E.Suburbs	3752	380	4736	0	0	0	0	0	4	0	8871
S.Suburbs	464	1102	2253	84	0	0	0	0	0	0	3903
Wgtn CBD	683	62	563	29	0	60	92	0	0	107	1597
J'ville	98	193	1523	443	0	0	0	0	0	43	2299
Tawa	0	0	1154	0	0	26	0	0	0	0	1180
Porirua	36	0	707	5	157	826	152	36	0	0	1917
Mana	0	31	1012	0	43	35	216	157	0	0	1493
Kapiti	0	0	654	0	0	45	0	527	0	0	1226
Upper Hutt	0	0	851	6	0	0	0	0	1375	251	2483
Lower Hutt	0	454	4660	31	0	0	0	0	372	3809	9327
Total	5032	2221	18111	599	200	992	460	719	1751	4210	34296

The fit of the assigned matrices before and after adjustment has been improved so that observed versus estimated patronage is almost identical, as shown in Figures 3.10 and 3.11.

Figure 3.10 Plot of observed 1996 versus assigned 1988 public transport counts.

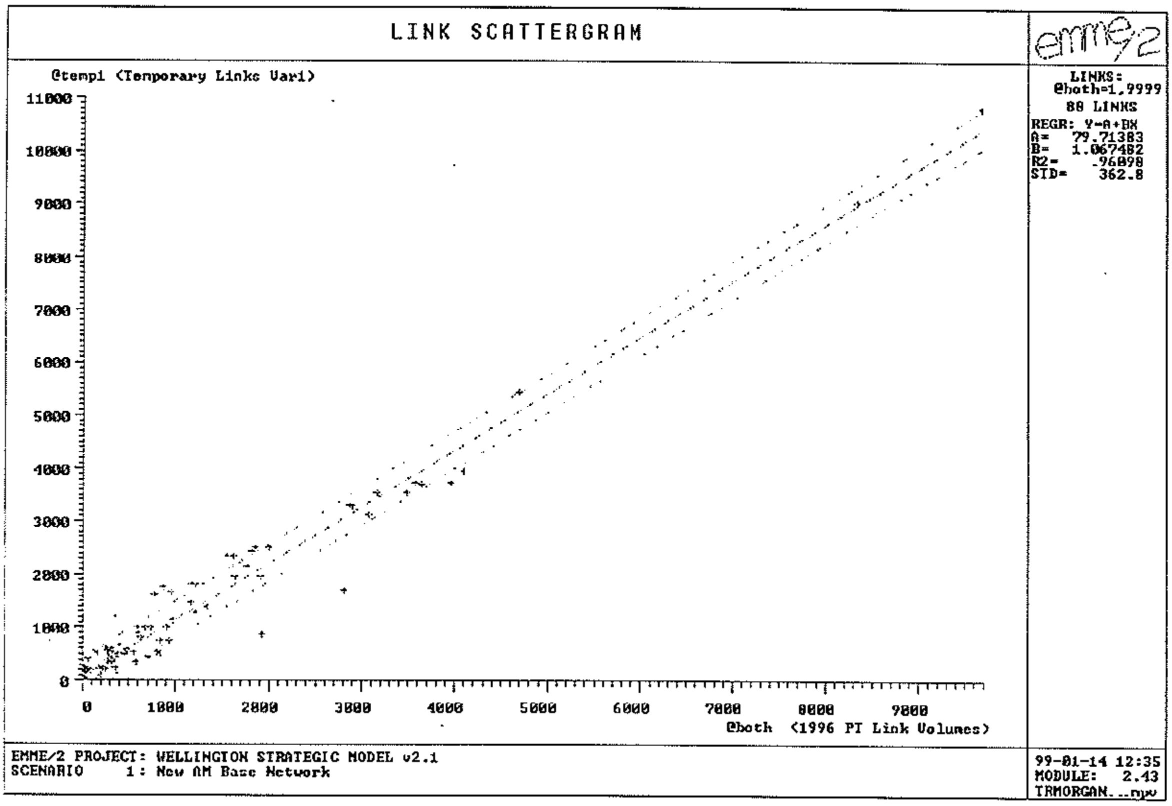
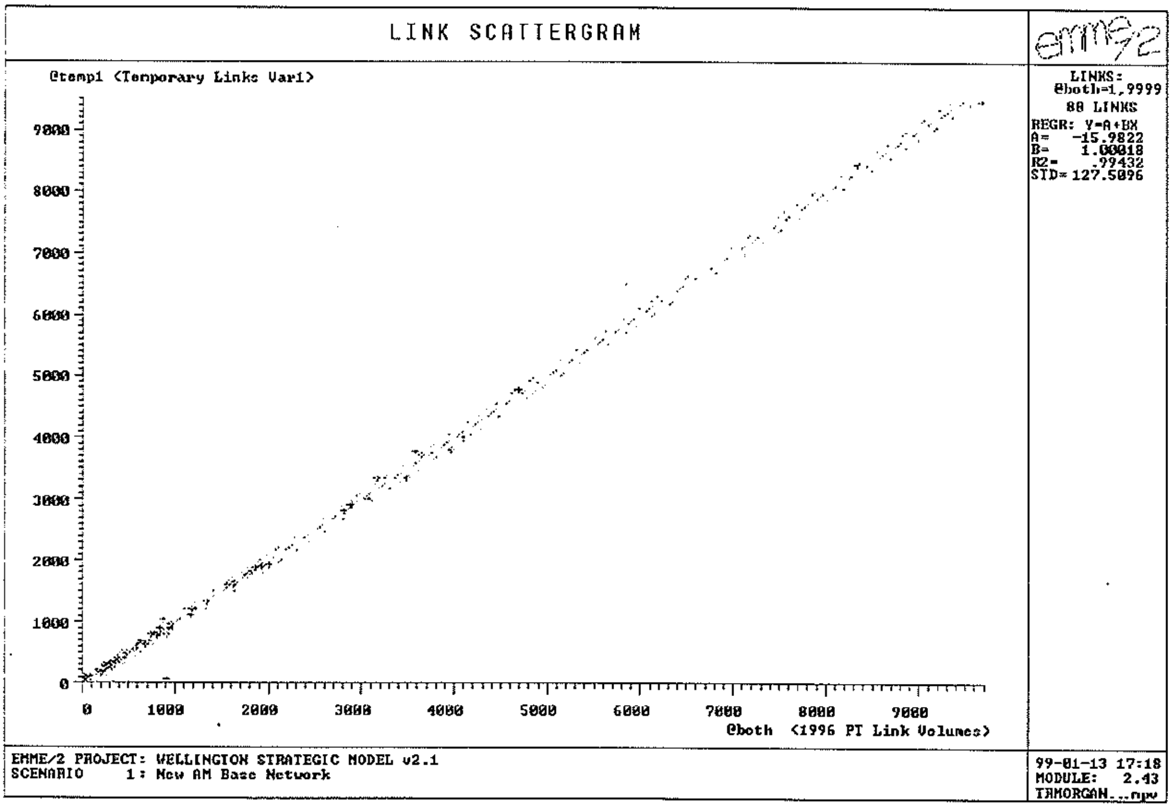


Figure 3.11 Plot of observed 1996 versus assigned 1996 estimated public transport counts.



After 100 iterations, convergence appears to have been reached (Figure 3.12), albeit slower than when the two (rail + bus) modes are considered independently (Figures 3.6 and 3.9). This is related to the interaction between rail and bus during the transit assignment.

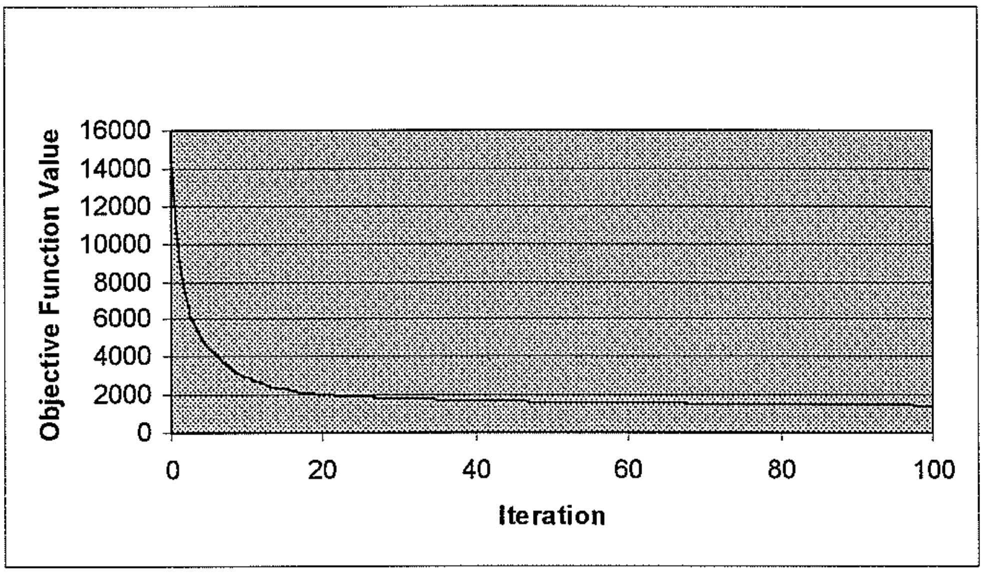


Figure 3.12 Objective function convergence for estimation of public transport demand.

4. CONCLUSIONS AND IMPLICATIONS

4.1 Approaches Available

A number of methods of varying complexity are now available for estimating public transport matrices. All the approaches studied for this project used variations on the gradient method to optimise an objective function, subject to constraints.

The objective functions vary between methods, with the three principal objective functions that are used being entropy maximisation, least squares, and maximum likelihood. The maximum likelihood formulation is statistically superior, but computationally more complex. The method adopted for this project was to minimise the sum-of-squares of the differences between assigned and observed link volumes. The algorithm was tested using the Wellington regional transport model and data. This is an EMME/2 model and the algorithm was written in the EMME/2 macro language. However the resulting algorithm could be implemented using any standard public transport assignment program, and could probably be implemented manually for a small radial network.

4.2 Testing the Effectiveness of Chosen Approach

The methodology by itself has limitations when applied to a seed matrix which varies significantly in size or shape from the observed matrix. The main problem is that only cells for which there is information are included in the optimisation. O/D flows which do not cross screenlines are not updated. The procedure can be improved by adjusting the seed matrix before applying the algorithm to take account of other known information. This prior information could include:

- Total trip numbers,
- Total revenue,
- Total arrivals and departures by zone,
- Trip length distribution,
- Changes to trip origins related to new housing developments, etc.

The cells which have not been updated can be isolated, and be adjusted after applying the algorithm. This procedure would be appropriate if, for example, the seed matrix was an historical data set or a sample data set, and it would be reasonable to calculate the average change in the cells which have been updated. The same average change could be applied to the cells for which there was no information.

These prior- and post-adjustments help address another shortcoming of the approach, which is a failure to utilise all the information available.

The method was shown to give plausible results when used to update the 1988 rail and bus matrices to 1996 observed data. These tests were undertaken for each mode individually without restricting the modes available. This provided a cross check on the model itself.

The algorithm converges monotonically towards a limit. The limiting value appeared to be non-zero but it was small in each case. This could arise because the assignment does not accurately replicate passenger behaviour, but is inevitable with models of this kind which represent PT services in terms of average frequency, etc. The size of the discrepancy is well within the normal error bounds of the model.

The most important conclusion is that the performance of the approach is highly dependent upon the quality of the seed matrix.

4.3 Implication for Data Collection

The importance of ensuring that all significant flows pass through cordons has implications for the choice of locations used for the cordons for quarterly transport surveys.

It is possible to separate out the unmodified component of the seed matrix. This could be assigned to the network, and locations requiring additional cordon points could then be identified.

4.4 Topics for Further Work

A number of topics have been identified where further work would be appropriate:

- Use of link versus line data,
- Use of station boarding and alighting data,
- Use of trip length distribution data to modify seed matrix,
- Solving small PT network estimations,
- Independent data checking,
- Weighting the data to reflect reliability,
- Analysis of trade-off between data collection costs and accuracy.

Use of link versus line data
The algorithm adopted could have been applied to counts obtained from each route crossing the cordon individually, rather than to those from the total cordon count. The latter approach was adopted however because limitations in the data, e.g. the incidence of missing or incomplete counts, affected some routes more than others.

By aggregating the data, some information was lost. Further analysis would be desirable to determine the appropriate trade-off in this case.

Use of station boarding and alighting data
The rail data were converted for this project from station boarding and alighting data to link counts. Again this resulted in some loss of information. Further work may be justified to either modify the PT network or to capture this information in another way. EMME/2 calculates boardings and alightings, but does not store this information in a way which would enable comparisons to be made. Otherwise these data could be used instead of the link data. This is a feature of EMME/2 which may not apply to other models.

Use of trip length distribution data to modify seed matrix
The importance of the seed matrix to the estimation process has been noted. While the literature identifies a general optimisation problem which effectively treats all PT-demand data as equal, the problem has had to be divided into two steps: pre-optimisation and optimisation. This concept could be refined further to incorporate trip length distribution data from ticket sales, to modify the seed matrix.

Solving small PT network estimations
The algorithm uses easily computable data which should enable a small problem to be solved using a spreadsheet model. However a spreadsheet model has not been tested in this project.

Independent data checking
In the matrix adjustment runs reported in Section 3, all the link data that were available have been used, and therefore no independent data were available for validation. The comparisons shown, therefore, tend to give an optimistic view of the accuracy of the results. Validation of the data from independent sources would increase the confidence in the performance of the algorithm.

Weighting the data to reflect reliability
Survey data from several sources are often in partial conflict. Where this is the case, it would be desirable for the method to provide for weighting of the data to reflect its judged reliability. Such weighting is incorporated in the general objective function (Section 2.2.1 of this report), but could be included with relative ease.

Analysis of trade-off between data collection costs and accuracy
Increasing the sophistication of the approach is likely to give decreasing returns in terms of accuracy versus cost of data collection. Further work on the sensitivity of the algorithm would enable this trade-off to be established.

5. REFERENCES

Heydecker, B.G., Verlander, N.Q., Bolland, J.D., Arezki, Y. 1995. Improved matrix estimation under congested conditions. *Traffic Engineering and Control 36(12):* 667-672.

MVA (Martin Vorhees Associates). 1997. Background Paper: Matrix Estimation Methods. January 1997.

Spiess, H. 1987. A maximum likelihood model for estimating origin-destination matrices. *Transportation Research 21B(5):* 395-412.

Van Vliet, D., Willumsen, L.G. 1983. Validation of the ME2 model for estimating trip matrices from traffic counts. *Proceedings of the Eighth International Symposium on Transportation and Traffic Theory:* 640-655. Toronto, Canada, June 1981.

Willumsen, L.G. 1984. Estimating time-dependant trip matrices from traffic counts. *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory:* 397-411. Delft, The Netherlands, July 1984.

AN IMPORTANT NOTE FOR THE READER

The research detailed in this report was commissioned by Transfund New Zealand.

Transfund New Zealand is a Crown entity established under the Transit New Zealand Act 1989. Its principal objective is to allocate resources to achieve a safe and efficient roading system. Each year, Transfund New Zealand invests a portion of its funds on research that contributes to this objective.

2. DEVELOPING THE METHODOLOGY

2.1 Background

The need for good origin–destination (O/D) data, and the high cost of undertaking the surveys to obtain such data, has led to the development of a variety of techniques for creating matrices[1] for estimating the PT-demand that are based on link data (Van Vliet & Willumsen 1983; Spiess 1987). Link data are much cheaper to obtain, particularly for highway studies where it can be collected by automated counters or manual counts for particular vehicle types.

Most early work on matrix estimation concentrated on solving the O/D data problem for matrices used for highways (Willumsen 1984), although the techniques have since been applied to public transport matrices. Even so, until recently the computation time and data storage requirements of the methods limited their use to very small problems. Increasing computer speed and hierarchical estimation techniques (MVA 1997) now enable very large matrices to be estimated. Passenger trip-matrix estimation approaches have also been developed (e.g. approaches described by MVA 1997, and Spiess 1990[2]), the basic principles being identical to the highway demand case.

2.2 Types of Approach

A review of recent literature carried out for this project indicated that most matrix estimation techniques are variations on the same basic approach. This is to use traffic counts as the primary input, together with a 'seed' matrix, which might be a matrix from a previous survey for example.

The early approaches formulated the problem as an optimisation problem in which the objective was to minimise the difference between the estimated demand matrix and the seed matrix. This approach was subject to the constraint that the assigned[3] volumes from the estimated[3] matrix should correspond to the observed volumes on links with count data.

The method assumed that errors lay entirely in the prior estimate of the matrix, rather than having a component in the traffic counts, and therefore it could be unreliable when errors were from another source.

---

1 A public transport trip matrix $\boldsymbol{M}$ is defined, with the origin and destination for a trip represented by the rows and columns respectively. Values in each cell m_{ij} gives the number of trips that are being made from origin zone '*i*' to destination zone '*j*' for a given time period. Typically these matrices are square (i.e. have the same numbers of rows as columns).

2 Information (dated May 1990) on a gradient approach for the O-D matrix adjustment problem by H. Spiess was obtained via an Internet communication.

3 Assigning - technique used to convert data estimated for a trip matrix to link data.
Estimated - data estimated in a trip matrix that represents demand for a PT trip.

2.2.1 General Approaches

More recent approaches have also formulated the problem as an optimisation problem, but used as an objective function the minimisation of the difference between all observed and modelled values. The inclusion of all data in the objective function allows a trade-off between conflicting sources of information (Heydecker et al. 1995).

Following Heydecker et al. (1995), a general form of the objective function is:

$$Z(F, f, w) = \sum_{b \in B} w_b \, S\,(F_b, f_b) \qquad (1)$$

where:

F	is a vector of model values,
f	is a vector of corresponding data or prior estimates,
w	is a vector of importance weights for the data,
$S\,(F, f)$	is a numerical measure of similarity between F and f,
B	is an index set of all data used, and
b	is an observation.

The model values F can be expressed in terms of a matrix m_{ij} by the relationship:

$$F_b = \sum_{od} Q^{b}_{od} \, m_{ij} \qquad (2)$$

where:

Q^{b}_{od}	is the proportion of traffic from o (origin) to d (destination) present in observation b.

There are three main forms of the measure of similarity $S\,(F,f)$:

- Entropy,
- Maximum likelihood,
- Least squares.

Each has advantages and disadvantages. Maximum likelihood theory proves that, if this form of objective function can be maximised, then the solution is statistically most consistent with the input information. However it requires a highly sophisticated optimiser, and there is no guarantee that the solution found is a global (rather than a local) optimum.

The maximum entropy and the maximum likelihood approaches minimise the sum of the squares of the relative error, while the least squares approach minimises the sum of the squares of the absolute errors. The least squares approach was chosen for testing in this study because, computationally, it is simpler to implement.

2.2.2 Gradient Approach

All the methods studied solved the optimisation problem using some variation of the gradient approach[2] or method of steepest descent. The method requires a seed matrix, which is assumed to be an estimate, to be adjusted as part of the process.

The gradient approach works by calculating the change in the objective function resulting from a small change in the estimated demand matrix – this is called the gradient matrix – and subtracting a multiple of the gradient matrix from the previous estimated matrix. The multiple is called the 'step length', and is chosen to maximise the reduction in the objective function at each step.

A large number of possible O/D matrices exist that could give rise to a particular observed set of link flows. The gradient method attempts to replicate the observed data with the minimum change to the seed matrix. This is both a strength and a weakness. The algorithm used only adjusts flows for which observed data are available. Also it adjusts only non-zero matrix elements. This has significant implications for the choice of both the seed matrix and the screenline points.

The algorithm works iteratively, with the resultant matrix from each iteration being used as the seed matrix for the next.

2.3 Proposed Method

The method selected for testing in this project defines the objective as being to minimise the sum of the squares of the differences between the observed passenger counts and the assigned passenger counts on selected links. The optimisation problem is solved using a gradient method.

The choice of approach was made on the basis of computational tractability. While more sophisticated methods are available, these require expensive proprietary software. The approach that has been adopted for this project can be implemented with a simple algorithm, which could also be applied manually to small networks.

Because this approach only uses the link data for optimisation, the optimisation function is a special case of the general function (Equation 1), where the weights of all other data are set to zero. However, a feature of the gradient method is that the approach produces a solution which minimises the change from the seed matrix.

As for all gradient methods, the approach requires two types of data:

- Matrix data (i.e. seed matrix),
- Link data (i.e. counts).

The link data are assumed to be the 'correct' data, which the algorithm will attempt to replicate by adjusting the seed matrix. They could be count data collected at selected screenlines or derived from boarding and alighting data.

The problem is thus to minimise $Z(m)$ defined as:

$$Z(m) = \sum (v_a - \bar{v}_a)^2 \qquad (3)$$

where: $\bar{v}_a$ is the observed count on the link a, and v_a is the volume count on the link resulting from the assignment of demand matrix m.

The problem is solved iteratively, with the demand matrix at iteration $n+1$ given by:

$$m_{i,j}^{n+1} = m_{i,j}^{n} \times \left(1 - \lambda^{n} \left[\frac{\partial Z(m)}{\partial m_{i,j}} \right]_{m_{i,j}^{n}} \right) \qquad (4)$$

$$G_{i,j} = \left[\frac{\partial Z(m)}{\partial m_{i,j}} \right] \text{ is the gradient matrix} \qquad (5)$$

and λ^n is the step length for sample n.

The gradient matrix represents the change in the objective function resulting from a small change in the estimated demand matrix. Spiess (1990) has shown that the gradient matrix is the matrix calculated by summing the differences between the observed and the estimated counts along each O/D path, i.e.

$$\frac{\partial Z(m)}{\partial m_{i,j}} = \sum_{k \in K_{i,j}} P_k \sum_{a \in A} \delta_{a,k} (v_a - \bar{v}_a) \qquad (6)$$

where:

- K_{ij} is the set of valid paths from i to j,
- P_k is the probability of using path k,
- $\delta_{a,k}$ is 1 if a is contained in path k, and 0 otherwise, and

The optimum value of λ (the 'step length') can be shown to be a function of the differences between the observed and assigned trips and the link 'volumes' that result from assigning the gradient matrix (G).

$$\lambda^* = \frac{\sum (assign(G) \times (v_a - \bar{v}_a))}{\sum (assign(G))^2} \qquad (7)$$

where:

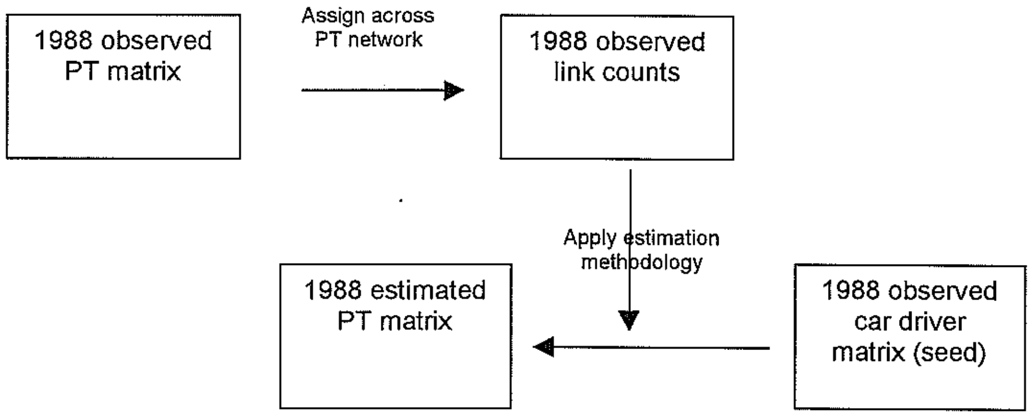
- λ^* is the optimum step length, and
- *assign (G)* is the 'volume' on link a resulting from the assignment of the gradient matrix G.

Intuitively, because each element of the gradient matrix is the sum of the 'errors' on the links between each origin and destination, the gradient matrix has the desired properties (the O/D paths which cross the screenlines having the largest errors will be adjusted most), but the 'errors' will be counted a multiple number of times.

The optimum step length λ^* can be seen to factor the gradient matrix in proportion to a weighted sum of the errors divided by a weighted sum of the result of assigning the gradient matrix.

3.2 Validation of the Methodology

The validation runs were undertaken using link data from the WRC model created by assigning the 'known' observed matrix, and using a different matrix as the 'seed'. The following is a diagrammatic representation of the process undertaken:



Three different seed matrices were used: car, rail, bus; and then a combination of two sets of data (rail + bus).

3.2.1 1988 Car Seed Matrix

The first matrix tried was the 1988 observed car driver matrix as a seed. Use of the car matrix without any pre-adjustment demonstrated the potential dangers of the approach. The estimated PT matrix 'fitted' the observed data points on the network well, but some cells in the estimated matrix were unchanged by the process. This is a desirable property when the seed matrix has a similar size and shape to the matrix being estimated, but in this case it left values that were clearly erroneous.

The process was then modified by factoring the car driver matrix to give the correct total number of trips. The results of this experiment are shown in Tables 3.1 and 3.2 which are the initial observed car seed and PT matrices, and in Table 3.3 which is the result of applying the algorithm to the scaled matrix.

Comparison of the three matrices indicates that, while the shape of the car driver seed matrix was far removed from what was being estimated, the methodology updates this seed data into an estimate which is not too dissimilar.

Although the seed matrix was modified to be the correct size, the resulting matrix is still too large. A closer result was obtained by further manipulation of the final matrix.

By taking the difference between the seed matrix and the estimated matrix, matrix cells that have changed under the process can be identified. The cells that have changed are the best estimates for the PT matrix. Cells which have not changed are either O/Ds not connected by public transport, or where there is no updating information.

Figure 3.1 Plot of 'observed' versus 'estimated' matrices before modification.

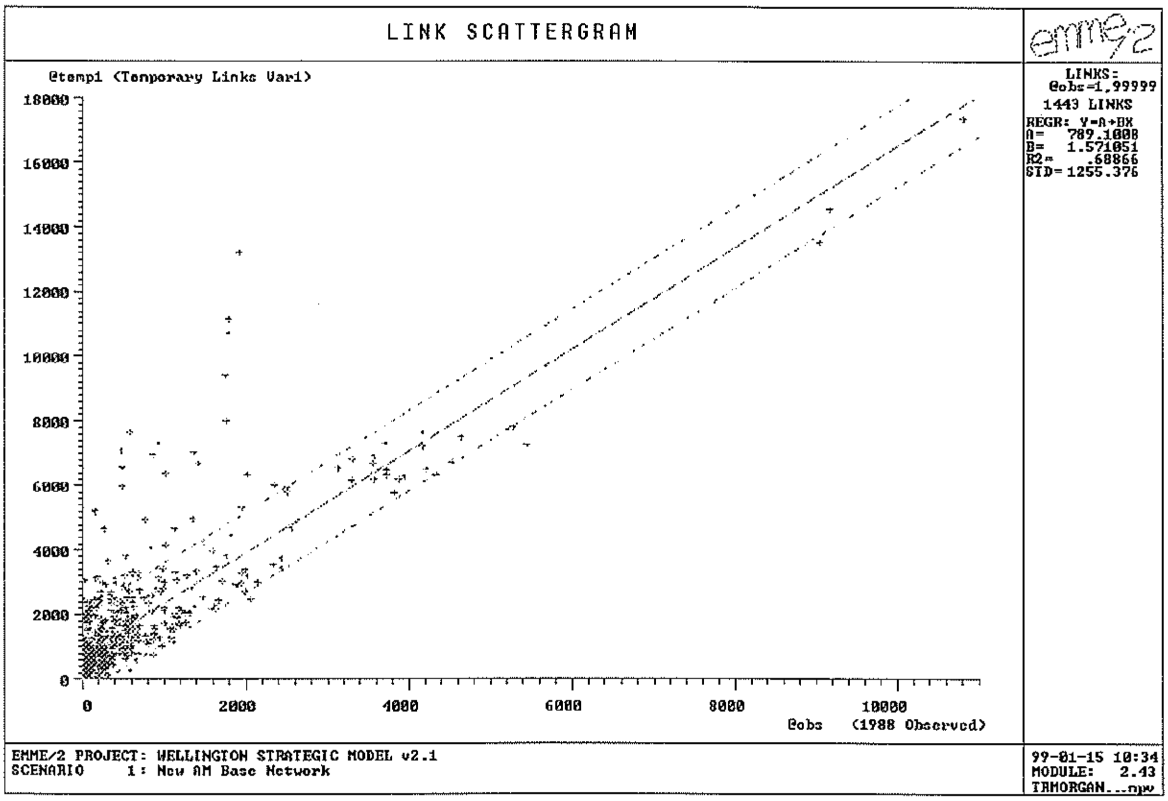
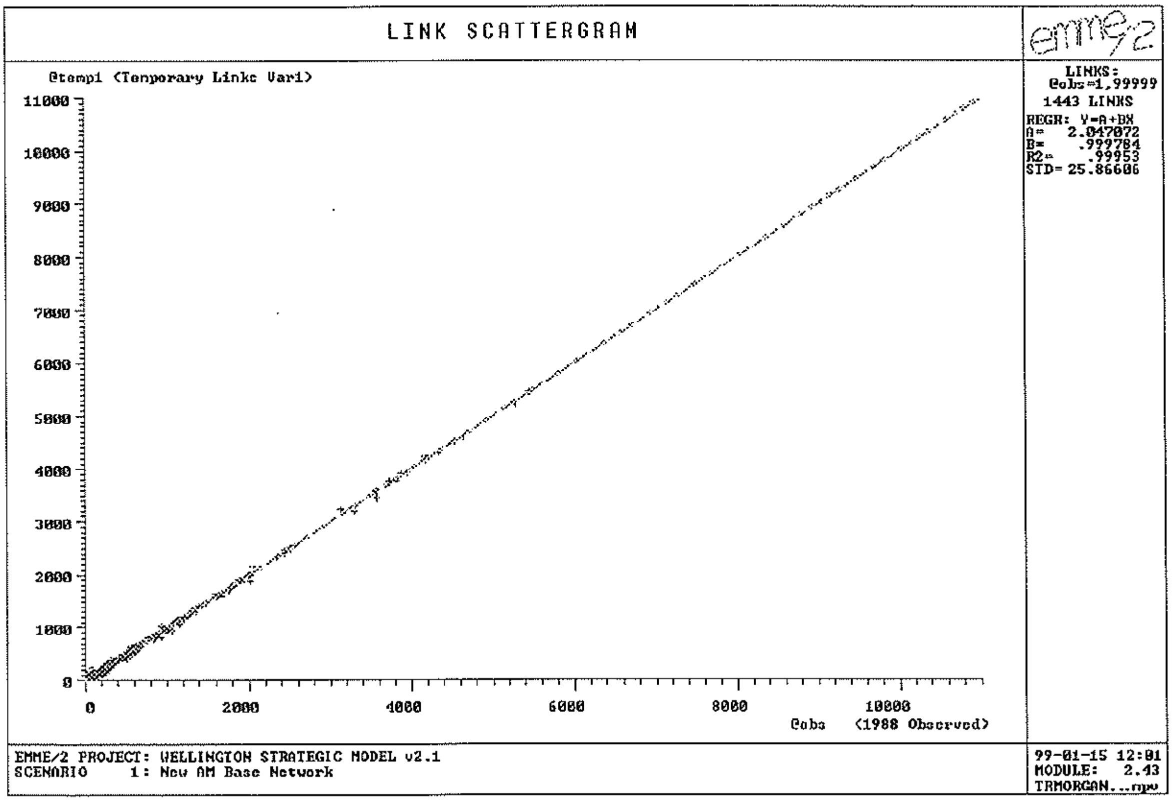


Figure 3.2 Plot of 'observed' versus 'estimated' matrices after modification.



The algorithm was set to run for 300 iterations, and the value of the objective function (the sum of the squares of the errors) after each iteration was plotted. This is shown in Figure 3.3.

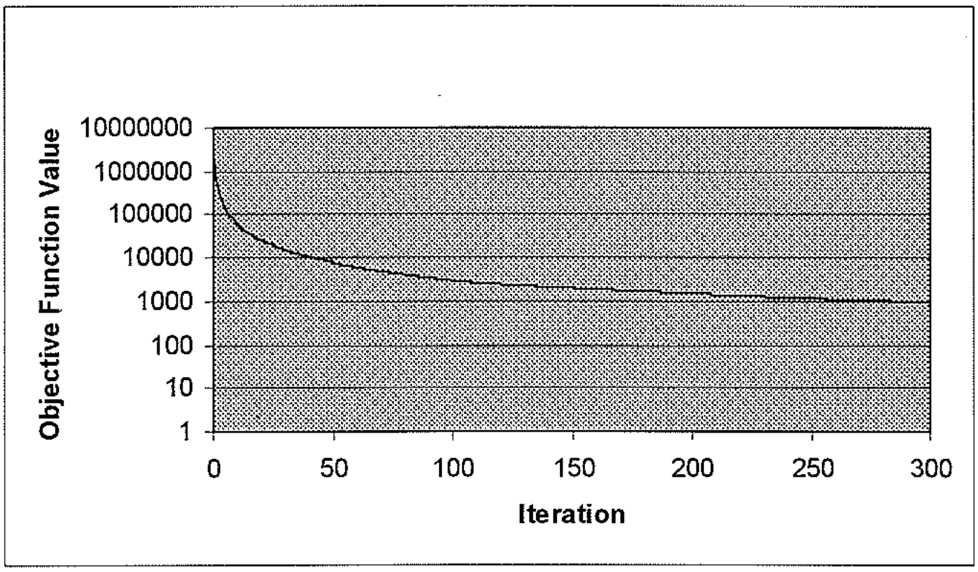


Figure 3.3 Objective function convergence for testing methodology.

As can be seen, the value of the objective function falls rapidly initially and is monotonic (i.e. does not oscillate). These are highly desirable properties. Note that the scale of the y-axis is logarithmic, and that the value of the objective function does not actually reach zero. This results from a combination of interaction between rail and bus during the transit assignment, and anomalies within the specified PT network within the Wellington model.

3.2.2 1988 Rail Seed Matrix

The methodology was applied to the rail services using the 1988 observed trip matrix that was derived from the 1988 HIS, and updating it with count data collected by the WRC on census day in 1996.

Table 3.5 shows the observed morning (AM) peak rail trip matrix generated from the 1988 HIS, while Table 3.6 shows the resulting estimated 1996 trip matrix for the morning peak.

Comparison of the two matrices shows that, as expected, rail patronage decreased over the 1988-96 period. Rail trips from Johnsonville have reduced from 1,902 in 1988 to an estimated 1,181 (38% reduction) in 1996. In contrast, rail trips from the Kapiti Coast has increased by almost 70% over the same period, in keeping with the high population growth that has been observed in that part of Wellington Region.

Figure 3.4 Plot of observed 1996 versus assigned 1988 rail trip counts.

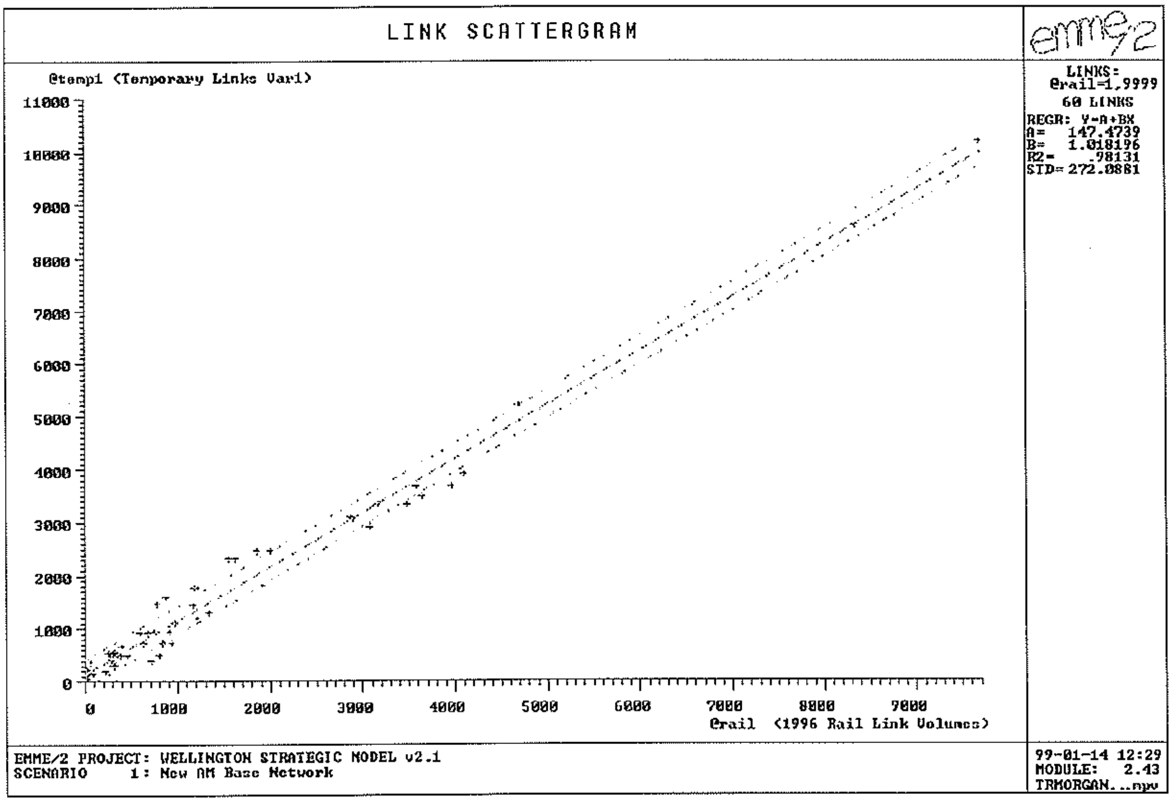
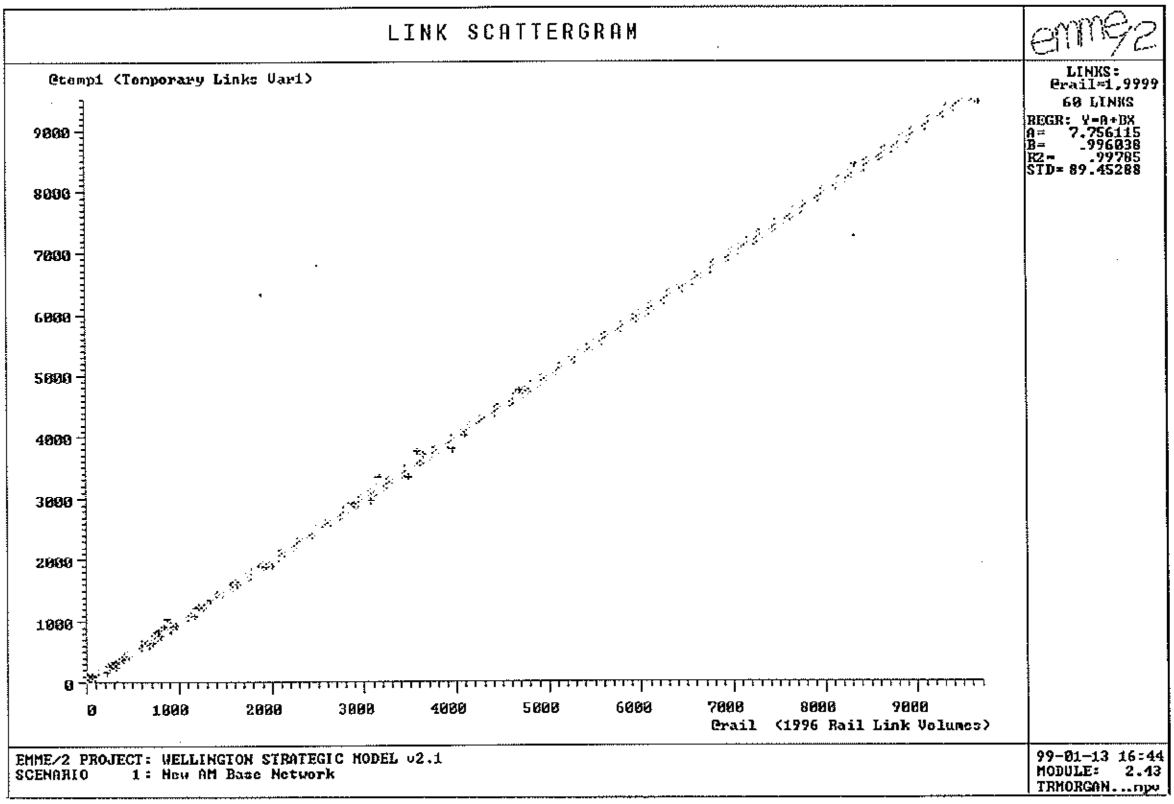
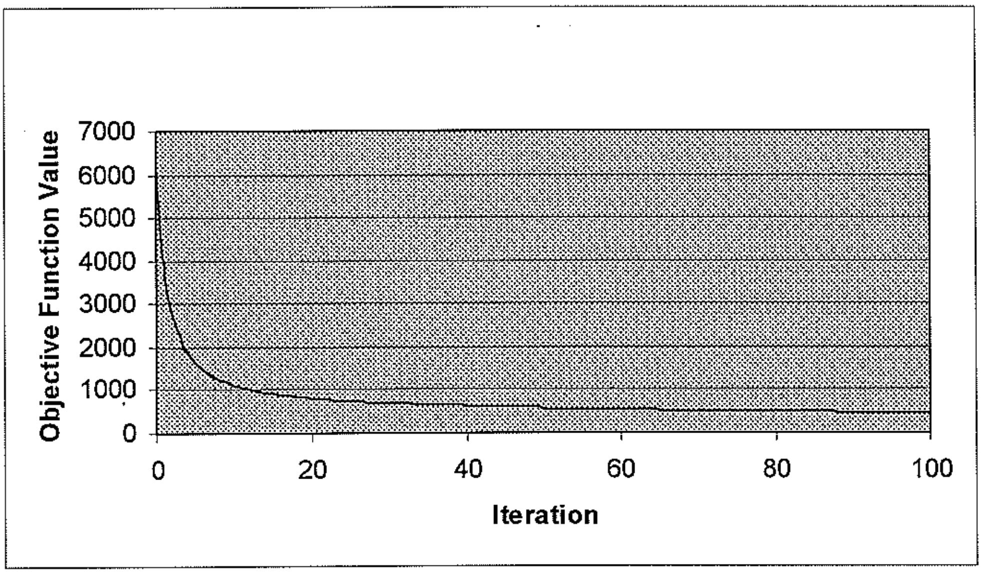


Figure 3.5 Plot of observed 1996 versus assigned 1996 estimated rail trip counts.



The algorithm was set to run for 100 iterations and the value of the objective function (the sum of the squares of the errors) after each iteration was plotted. This plot is shown in Figure 3.6.

Figure 3.6 Objective function convergence for estimation of rail matrix.



As can be seen, the value of the objective function falls rapidly initially and is monotonic (i.e. does not oscillate). These are highly desirable properties. The value does not actually reach zero. This results from the specification of the network. The residual is equivalent to 0.3 persons per cell of the matrix.

3.2.3 1988 Bus Seed Matrix

The methodology was then applied to the observed 1988 bus patronage matrix (Table 3.7). A 1996 patronage matrix was estimated using the 1996 census data (Table 3.8). The total patronage does not appear to have significantly changed over the 1988-96 period. Significant effects are an estimated increase in total bus trips from Lower Hutt (29%), and a decrease in bus trips to the Wellington CBD.

A population increase within Miramar and Seatoun over this period has contributed to an increase in bus patronage within that part of Wellington Region.

Internal patronage within the Lower Hutt area has increased by 11% because of an increase in bus accessibility through the development of new commercial services covering Lower Hutt, Wainuiomata, and Eastbourne during the time frame (1988-1996) studied. Some of these services have been extended to Courtenay Place in Wellington CBD, and therefore their attractiveness of using the bus for trips to the city has increased.

Figure 3.7 Plot of observed 1996 versus assigned 1988 bus trip counts.

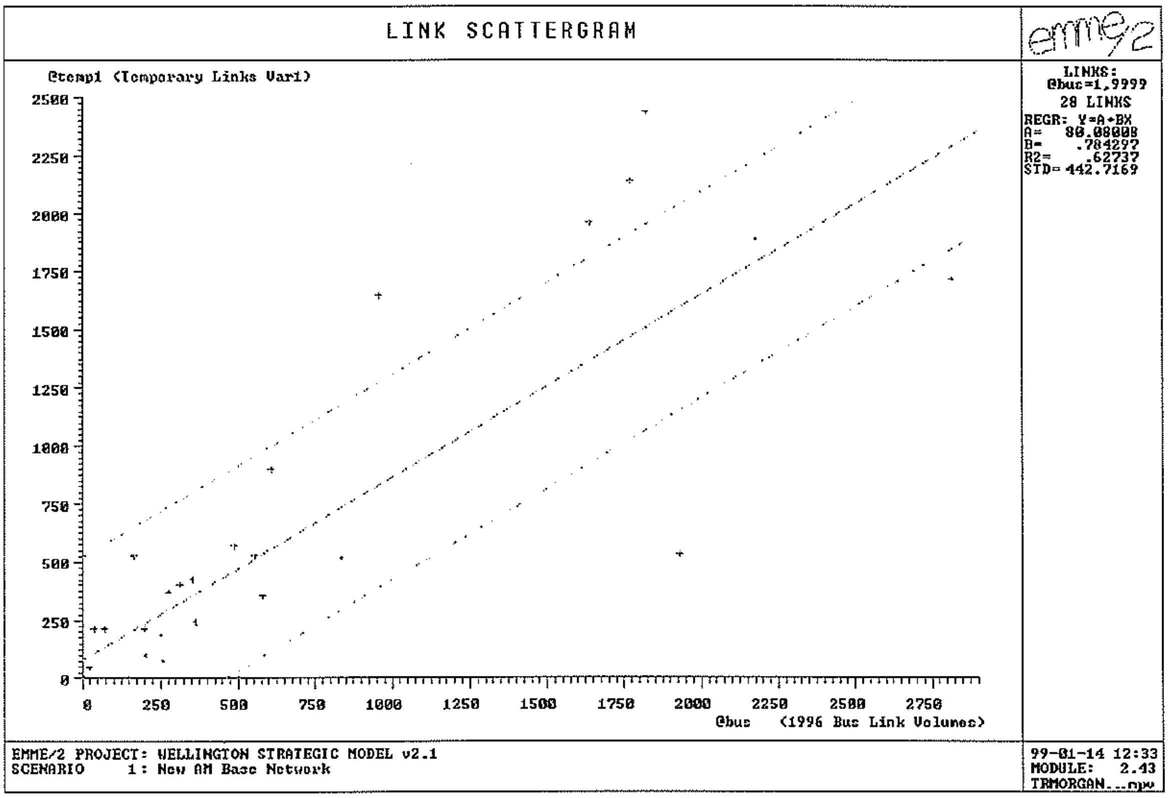
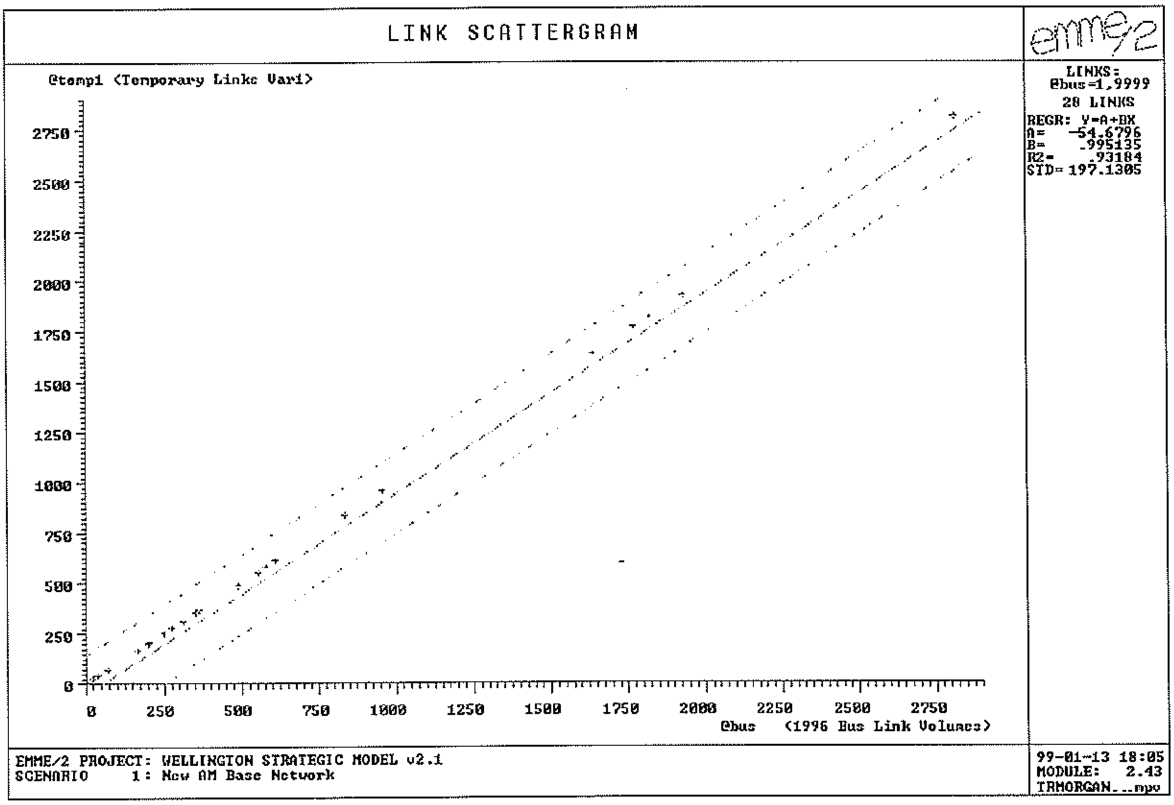
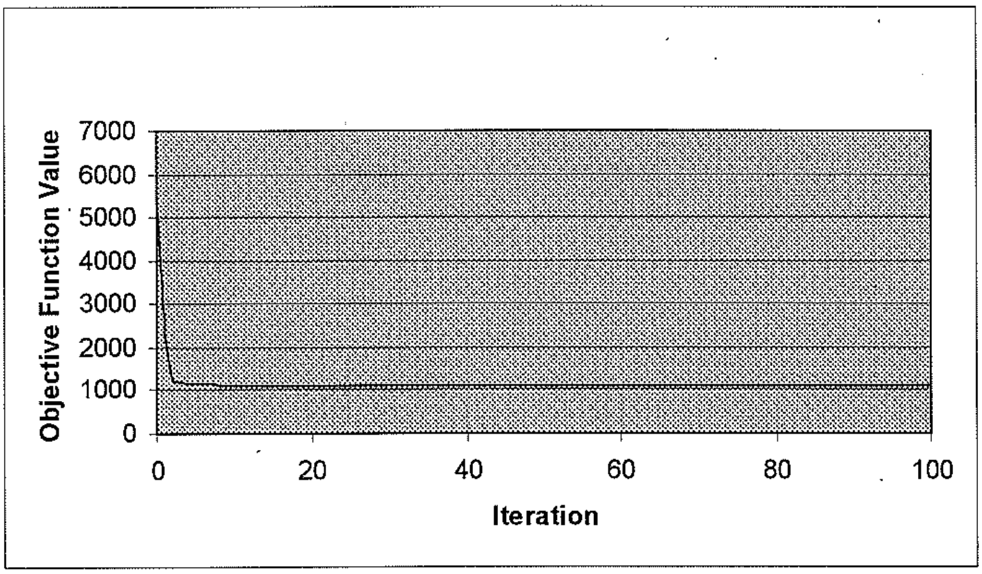


Figure 3.8 Plot of observed 1996 versus assigned 1996 estimated bus trip counts.



As previously, the convergence of the objective function was tested. The estimation was conducted over 100 iterations. Compared with the rail estimation, the bus convergence appears to occur more quickly. This is because of the nature of the bus network, where bus trips can take a choice of a number of routes, compared to the linear nature of the rail network.

Figure 3.9 Objective function convergence for bus trip estimation.



3.2.4 Bi-modal Estimation

The methodology was then tested on the combination of the two sets of data (rail + bus). The seed matrix used was the 1988 observed morning peak total patronage matrix. The total observed matrix (Table 3.9) is the sum of the observed rail and bus matrices. However the estimated matrix (Table 3.10) is not the sum of the rail and bus estimates because of the interaction within the transit assignment of EMME/2.

Overall, total public transport patronage is estimated to have decreased by 7% over the 1988-1996 period. The effects within the matrix are typically the sum of the effects that were estimated by the transit modes independently. The changes are an increase in total transit patronage from Kapiti (especially to the CBD), an increase in trips from Lower Hutt, a large decrease in trips from Johnsonville (especially to the CBD), and general decreases in other parts of Wellington Region.

Figure 3.10 Plot of observed 1996 versus assigned 1988 public transport counts.

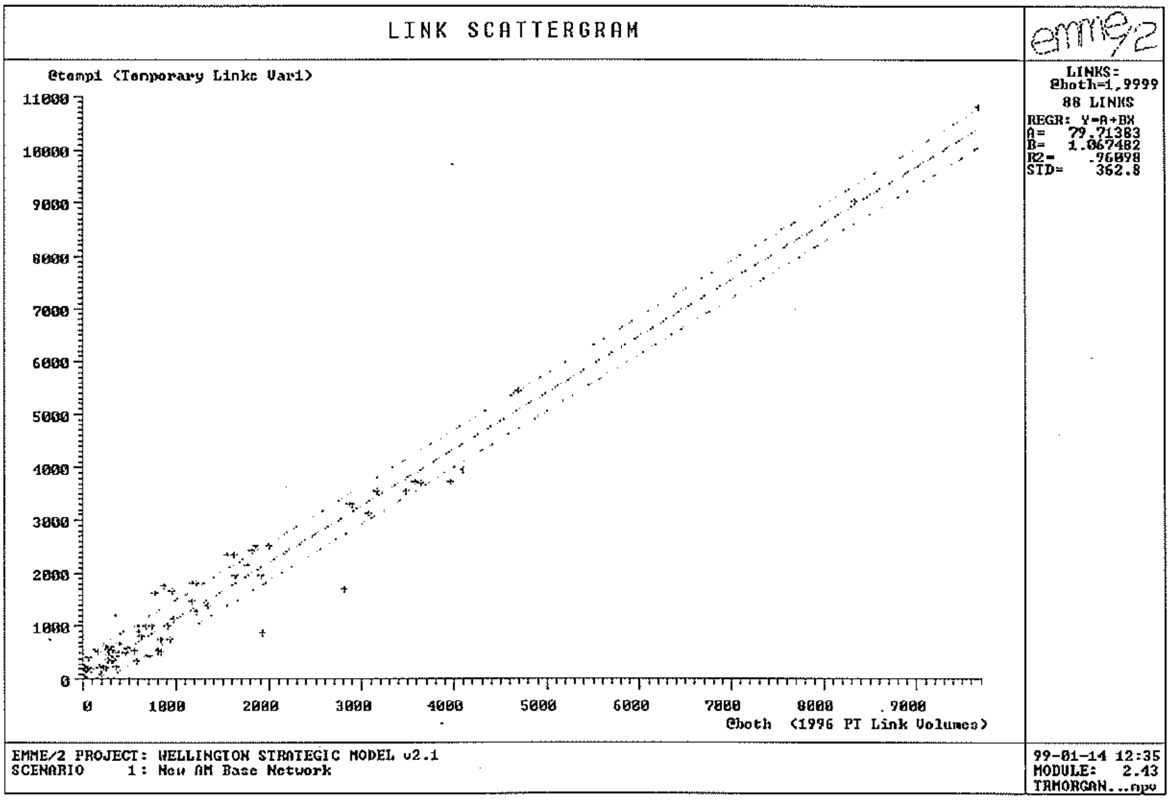
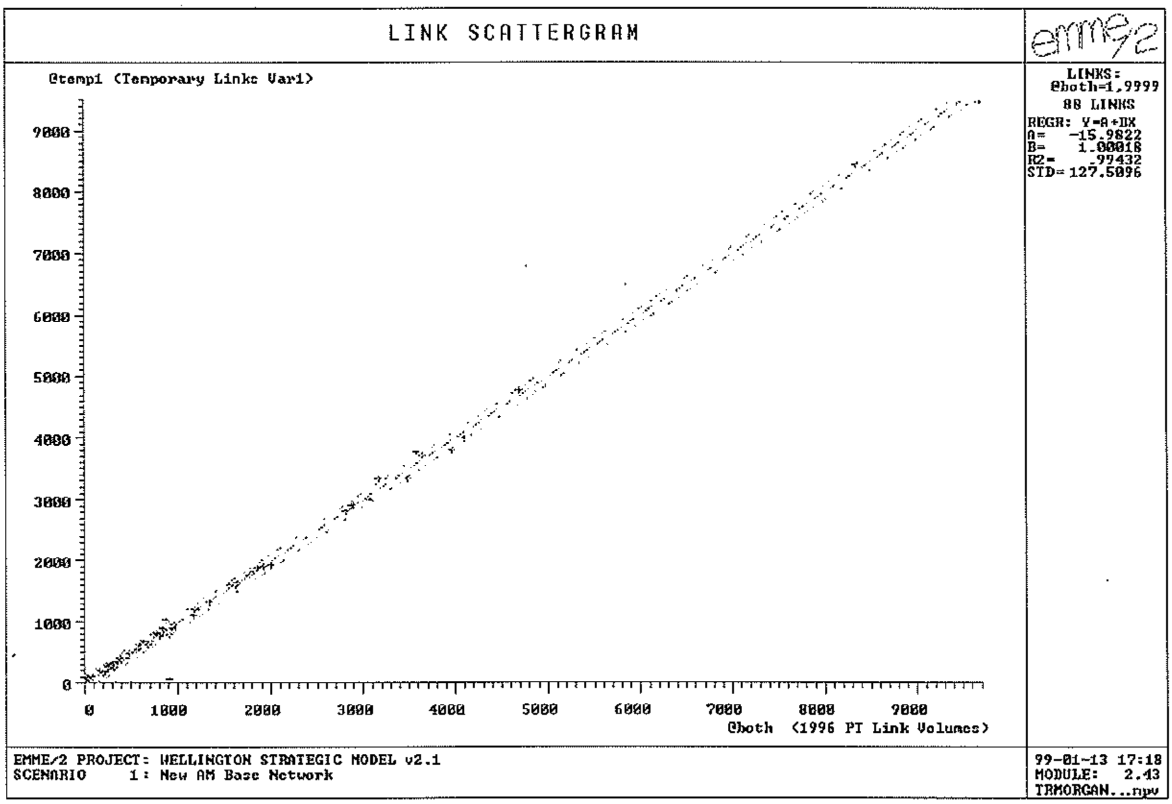


Figure 3.11 Plot of observed 1996 versus assigned 1996 estimated public transport counts.



After 100 iterations, convergence appears to have been reached (Figure 3.12), albeit slower than when the two (rail + bus) modes are considered independently (Figures 3.6 and 3.9). This is related to the interaction between rail and bus during the transit assignment.

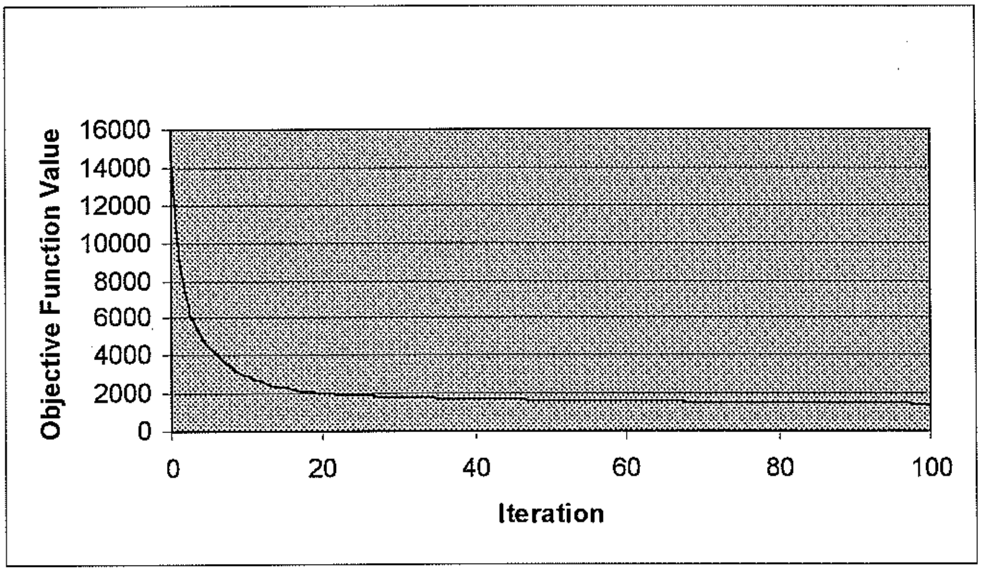


Figure 3.12 Objective function convergence for estimation of public transport demand.

5. REFERENCES

Heydecker, B.G., Verlander, N.Q., Bolland, J.D., Arezki, Y. 1995. Improved matrix estimation under congested conditions. *Traffic Engineering and Control 36(12):* 667-672.

MVA (Martin Vorhees Associates). 1997. Background Paper: Matrix Estimation Methods. January 1997.

Spiess, H. 1987. A maximum likelihood model for estimating origin-destination matrices. *Transportation Research 21B(5):* 395-412.

Van Vliet, D., Willumsen, L.G. 1983. Validation of the ME2 model for estimating trip matrices from traffic counts. *Proceedings of the Eighth International Symposium on Transportation and Traffic Theory:* 640-655. Toronto, Canada, June 1981.

Willumsen, L.G. 1984. Estimating time-dependant trip matrices from traffic counts. *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory:* 397-411. Delft, The Netherlands, July 1984.