DISTILLATE PROJECT F3c: DESIGN OF A SCENARIO INTERPRETER IN TRL'S STRATEGIC TRANSPORT MODEL

by Dr Andrew Ash

RPN 033

Final Version (May 2008) PROJECT REPORT

Transport Research Laboratory

PROJECT REPORT RPN 033

by Dr Andrew Ash
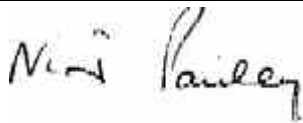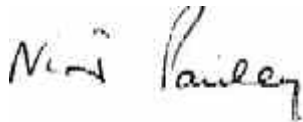
DISTILLATE PROJECT F3c: DESIGN OF A SCENARIO INTERPRETER IN TRL'S STRATEGIC TRANSPORT MODEL

CLIENT: EPSRC

	Name	Date Approved
Project Manager	*Nimi Pauley*	21.05.08
Technical Referee	*Nimi Pauley*	21.05.08

Contents

Executive summary

DISTILLATE (Design and Implementation Support Tools for Integrated Local Land use, Transport and the Environment), was a UK EPSRC (Engineering and Physical Sciences Research Council) funded project which sought to enable significant improvements in the ways in which sustainable transport and land use strategies are developed and delivered in the UK.

This report describes work on a case study carried out as part of Project F (Decision Support Tools) of the DISTILLATE project. The Local Authority partner for this study was Strathclyde Partnership for Transport (SPT).

The aim of the study was to develop a modelling technique which would assist interpretation of scenario outputs forecast by strategic transport models. In the present case the techniques have been applied to the Strathclyde application of TRL's Strategic Transport Model (STM). The technique would allow users to identify the likely drivers and mechanisms responsible for model outputs under particular land-use and transport policy assumptions The study can be seen as a response to the often-justified accusation that transport models are impenetrable 'black boxes' which generate results for which no easy explanation is at hand.

A key idea of this project is that the transport model should provide appropriate diagnostic outputs which are closely linked to the underlying mechanisms within the model, thus providing a good basis for interpretation. We see this approach as ultimately leading to an "intelligent interpreter" which will automatically construct a form of narrative account of the policy test outputs based on the model mechanisms. Within the present project we are necessarily limited to speculation in this direction, but the approach presented here should be a sound basis for future work.

This report has described a concept for a Scenario Interpreter and a prototype has been developed and demonstrated. Its purpose is to provide the user with an evidential basis for the interpretation of scenarios generated by the TRL Strategic Transport Model (STM). This will allow users to identify the likely drivers and mechanisms responsible for model outputs under particular land-use and transport policy assumptions.

We have identified the key techniques to be employed in the interpreter; central to the method is a technique of using a first order sensitivity analysis applied in the final iterative pass of the STM (as distinct from one based on simple re-running of the all model iterations). This has the advantage of reduced run times and a clearer relationship between outputs and model mechanism.

The report points to possibilities for further development of the software implementation of the model and we conclude that the technique is capable of considerable elaboration with automatic algorithms to carry out the analyses. We also look forward to a version of the interpreter which can provide results of analyses in natural language.

Abstract

DISTILLATE (Design and Implementation Support Tools for Integrated Local Land use, Transport and the Environment), was a UK EPSRC (Engineering and Physical Sciences Research Council) funded project which sought to enable significant improvements in the ways in which sustainable transport and land use strategies are developed and delivered in the UK.

This report describes work on a case study carried out as part of Project F (Decision Support Tools) of the DISTILLATE project. The Local Authority partner for this study was Strathclyde Partnership for Transport (SPT).

The aim of the study was to develop a modelling technique which would assist interpretation of scenario outputs forecast by strategic transport models. This report has described a concept for a Scenario Interpreter and a simple prototype has been developed and demonstrated. Its purpose is to provide the user with an evidential basis for the interpretation of scenarios generated by the TRL Strategic Transport Model (STM) but the general concept has wider application. The method described will allow users to identify the likely drivers and mechanisms responsible for model outputs under particular land-use and transport policy assumptions.

We have identified the key techniques to be employed in the interpreter; central to the method is a technique of using a first order sensitivity analysis applied in the final iterative pass of the STM (as distinct from one based on simple re-running of the all model iterations).   This has the advantage of reduced run times and a clearer relationship between outputs and model mechanism.

The report points to possibilities for further development of the software implementation of the model and we conclude that the technique is capable of considerable elaboration with automatic algorithms to carry out the analyses.

1   Introduction

DISTILLATE (Design and Implementation Support Tools for Integrated Local Land use, Transport and the Environment), was a UK EPSRC (Engineering and Physical Sciences Research Council) funded project which sought to enable significant improvements in the ways in which sustainable transport and land use strategies are developed and delivered in the UK.

This report describes work on a case study carried out as part of Project F (Decision Support Tools) of the DISTILLATE project. The Local Authority partner for this study was Strathclyde Partnership for Transport (SPT).

The aim of the study was to develop a modelling technique which would assist interpretation of scenario outputs forecast by strategic transport models. In the present case the techniques have been applied to the Strathclyde application of TRL's Strategic Transport Model (STM). The technique would allow users to identify the likely drivers and mechanisms responsible for model outputs under particular land-use and transport policy assumptions The study can be seen as a response to the often-justified accusation that transport models are impenetrable 'black boxes' which generate results for which no easy explanation is at hand. In consequence, users' confidence in models is undermined: they are left in doubt about the correctness of the software implementation and about the validity of the model assumptions. We recognise this to be a real problem and hope that the present work will be a contribution to a remedy for it. The use of a "scenario interpreter" is, we believe, especially desirable because every model has limitations to its validity or robustness and will inevitably produce "difficult" results which need interpreting. The use of an interpreter is not, of course, limited to or primarily with model pathology; it can be used to increase understanding of the functioning of the model in general.

A key idea of this project is that the transport model should provide appropriate diagnostic outputs which are closely linked to the underlying mechanisms within the model, thus providing a good basis for interpretation. We see this approach as ultimately leading to an "intelligent interpreter" which will automatically construct a form of narrative account of the policy test outputs based on the model mechanisms. Within the present project we are necessarily limited to speculation in this direction, but the approach presented here should be a sound basis for future work.

The main body of the report begins in Section 2 with an overview of various model variables and the way they are employed. The purpose here is to provide the reader with an understanding of model variables and mechanisms which form the basis of the Scenario Interpreter. Although written with STM in mind (detailed information on STM itself is provided in Appendix A) these considerations could apply to a wide range of strategic transport models. Section 3 presents the philosophy behind the scenario interpreter. Section 4 describes the model in detail and the underlying theory. In Section 5 an implementation of the Scenario Interpreter is described and outputs from the model are presented. Section 6 considers the further possibilities for the model and Section 7 completes the report with a summary and conclusions.

2 The strategic transport modelling process – basic concepts

2.1 Introduction

The following sections function as a kind lexicon for the rest of the report. Important concepts which underpin the STM (and other strategic transport models) are discussed in a form necessary for understanding the development of the Scenario Interpreter, described in later sections. It is also an opportunity to avoid later confusion by clarifying some commonly-used terms, if a little pedantically. Much, if not all of the following will be familiar to many but as it forms the logical springboard into later sections it may still be found of value to those who feel they know it already. Most of what is said is of general validity for strategic transport models based on demand models using the concept of generalised cost.

2.2 Spatial detail

Transport models operate on two levels of detail: the zonal level and the network level. At the network level a congestion submodel is used to predict the flows and journey times of persons or vehicles on the links of the road and public transport networks. At the zonal level the study area is divided into subareas (zones) and the model predicts the flows between zones. The zonal level is used for making predictions of trip generation, modal shares and trip distribution. At the network level an assignment model can be used to forecast the pattern of routes (and therefore the link loadings) arising from the demand for travel between zones. In the absence of an assignment model routes would need to be assumed and congestion model would calculate travel times on the basis of these.

A comprehensive model of travel demand would use both these levels, but at some computational cost. The network level would then be the source of travel costs ('skimmed costs') for the zone level demand model which, in its turn, provides the network level with zone to zone demand by mode. This somewhat simplifies a process which involves a number of technical problems associated with iterating between the two levels.

It should noted at this point that the version of STM (Version 4.6) used in this study does not use an assignment model and that its congestion calculations are based on a combination of link-type and area-wide speed/flow relationships (see Appendix A). Later versions (Version 5.0 onwards) can use a more conventional network based modelling approach.

2.3 Inputs

Model inputs mainly relate either to the forecast year or the base year. There may also be model parameters which are assumed to apply across all years. Base year data are dominated by matrices of trips (persons per hour), travel cost elements (travel times, fares etc) and planning data (population, jobs, retail floorspace etc) which are, where appropriate, broken down into various categories (e.g. trip purpose, household car

ownership). The forecast data are the planning data for the forecast year together with data defining the policies to be used in that year. In STM the policies are typically expressed as multipliers of base year cost components (e.g. fares and service levels) for travel between specific origin-destination pairs. Policies - along with congestion and public transport crowding effects – change the values of travel costs from their base year levels.

Some policies require additional forecast data. In STM, such policies relate to changes in the highway network and the introduction of both new modes (LRT and park and ride) and of new railway stations. It would take us too far to deal with these aspects of the model in any detail in the present report, but the principles set out in this report are also applicable to these cases.

2.4 Outputs

Outputs are the predictions of the transport model under assumptions about population (total number, age profile, and car ownership), land-use (residential distribution, jobs, retail floorspace) and transport costs (by private car and public transport and slow modes).

The term 'output' normally refers to what is written to a file or displayed on screen; however, one can generalise the term here to cover any end result of the modelling process whether it is physically output or not. For STM these outputs are

- The forecast trip matrices by mode, purpose, car ownership and time period
- The travel costs (e.g. travel times, parking costs, crowding costs)
- Factors dependent on the first two items (e.g. vehicle emissions)

STM trip matrix outputs are in fact mainly organised in terms of different aspects of the trip matrices (e.g. modal shares of trips to a zone). Not all costs information are outputs e.g. fares and service levels. The principal 'cost' output determined by the model is the highway travel times, which are controlled within the model by speed/flow relationships.

The outputs that this report is mainly concerned with are as follows:

- person trip attractions
- person trip starts
- area based person kms
- area based speeds
- area based car pcu kms
- emissions
- accessibilities to zones
- cordon counts by cordon area
- travel times to zones

2.5 Generalised costs

Travel choices depend on a range of factors, some tangible and measurable, others subjective and difficult to pin down mathematically. It has become normal practice in the UK to base choice models on travel costs or disutilities, known as "generalised costs". These costs combine, usually linearly, various money and time cost elements (mode specific) to form a single measure of cost. It is now a WebTag (Department for Transport) standard to measure these costs in units of time by converting money costs to an equivalent time by dividing the money cost by the value of time. More generally, weighting of different cost elements of the same type (e.g. time related) is also possible so as to reflect the trade-off rates between different elements of that type, e.g. valuing wait time more than in-vehicle time.

The STM requires base values for the generalised costs, which it constructs from exogenous data. When forecasting, it predicts a corresponding set of forecast year generalised costs (disaggregated by the same categories).

Transport policies can be expressed as multipliers of the base values for the appropriate cost elements e.g. a policy to raise bus fares by 10% in real terms is equivalent to multiplying base fares by 1.10.

2.6 Trip generation, mode shares and trip distribution

The fundamental objective of transport modelling at the zonal level is to predict the pattern of trip generation, modal share and trip distribution across the modelled zonal system. The link flows predicted by any network model depend on this. We therefore need to be clear about these terms. By 'trip generations' we mean the number of trips produced per hour within a zone; the trip rates would be defined as the average over a specified time period. Trip generations are usually broken down by categories which include purpose, time period and household car ownership, but other socioeconomic categories are also possible. The levels of home-based (HB) trips are driven by population and car ownership. Non-home-based (NHB) trip levels link activities at different locations where neither location is the residence of the traveller. NHB trip making is closely related to land-use factors such as job levels and retail floor space in the origin and destination zones.

Let the generations for origin o in time period t be denoted by gen(o,-, t). The '-' represents an unspecified list of categories (purpose etc.). Trips generated within a zone will terminate in destination zones. For each origin zone we can therefore define the proportion of trips generated which terminate in zone d in time period t as pod(o,d,-,t). This is the trip distribution function and it is the task of the trip distribution model to determine pod (o,d,-,t) in the forecast year. Changes in trip distribution are determined by changes in the costs of travel to alternative destination zones and changes in the distribution of land-use factors over those zones.

We can similarly consider modal shares. For each origin-destination zone pair the proportion of trips being made by a given mode can be denoted mod(o,d,-,t). The trip matrix can then be written as

$$T(o,d,-,t) = gen(o,-,t).pod(o,d,-,t).\mathrm{mod}(o,d,-,t)$$

From T(o,d,-,t) we can define column and row sums

$$T(o,*,-,t) = \sum_d T(o,d,-,t)$$

$$T(*,d,-,t) = \sum_o T(o,d,-,t)$$

which give the numbers of trips (per hour) starting and terminating in the origin and destination zones. These can be further summed over factors in the 'category list' such as purpose and household car ownership.

2.7 Choice models

In order to forecast modal shares and trip distribution shares given a base case it is necessary to use suitable 'choice models'. As an example of a model, consider a modal choice model. This predicts the modal shares for travellers between a given origin-destination zone pair from the base shares and the base and forecast generalised costs for each alternative mode.

The mathematical form for the choice models used can vary from model to model. Most commonly a logit model formulation is used:

$$p_i = \frac{p_i^0 \exp(-\lambda.\Delta c_i)}{\sum_k p_k^0.\exp(-\lambda.\Delta c_k)}$$

where the Δc are changes from the base in the generalised costs to each feasible alternative. The p_i are the forecast proportions and p_i^0 are the base year proportions. The equation applies to a given origin-destination zone pair and would also refer to particular trip/traveller categories (such as trip purpose) and time period.

The above formulation is an example of an incremental model: changes in shares are driven by cost differences so that if forecast costs and base costs are equal then the above equation automatically reproduces the base shares. This is to be contrasted with non-incremental approaches in which shares are forecast using only the forecast year costs (and depend therefore much more on the form of the model and its calibration).

The final outputs of a full and completed transport modelling process (after iterating to equilibrium and applying land-use effects) would only be guaranteed to match the base if forecast and base planning data were the identical.

In general, models use hierarchical arrangements of these multinomial logit models. The nesting of related alternatives (e.g. public transport) together allows a more realistic and flexible model structure. The model structures available within STM, which can be found more generally, are illustrated below in Figures 1 and 2.

These are the two hierarchies available in the latest version of STM (Version 5.2 – this is not used in this study but the models are a useful starting point) for a given origin, purpose and household car ownership. The hierarchy contains models for time period(t), mode(m) and distribution(d) and also Park and Ride (this should be ignored for present purposes). The diagrams suppress the presence of three nests for mode (car, slow modes, public transport). Where an alternative is represented by a lower level nest the generalised cost is a composite cost formed from the market shares and costs of alternatives within the nest.

In the current versions of the Strathclyde STM time choice is not activated and the model uses fixed proportions for the time period distribution (one should therefore imagine the time 'bubble' to be omitted.
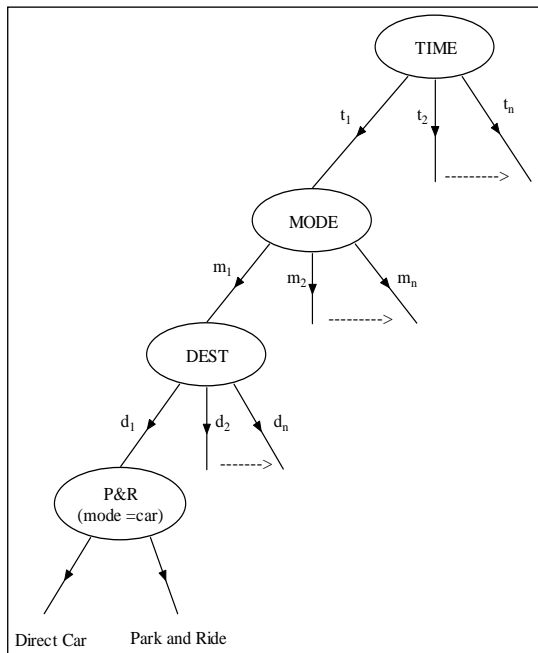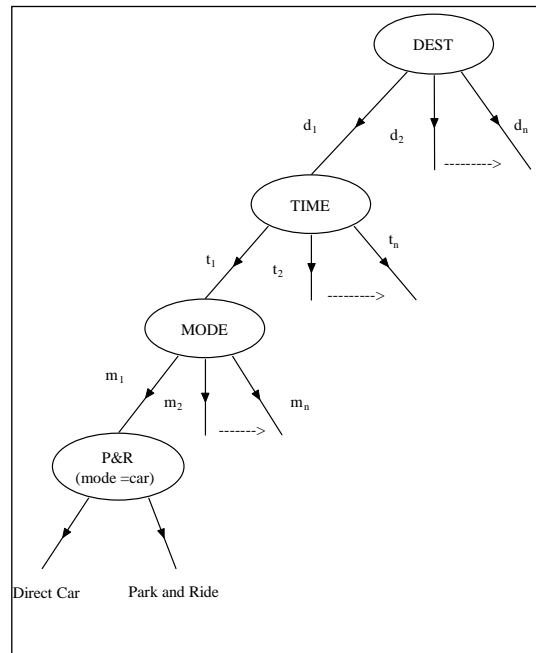

Figure 1:   t-m-d hierarchy   Figure 2: d-t-m hierarchy

The destination choice mechanisms also have constraints placed on the numbers of trips going to each destination based on the redistribution of land-use factors such as retail floor space and jobs.

The version of STM used in this study (Version 4.6) used a different, power law form for the modal and distribution choice model. The modal choice model can be written

$$p_i = \frac{p_i^0 \exp(-\lambda_2.\ln(c_i/c_i^0))}{\sum_k p_k^0.\exp(-\lambda_2.\ln(c_k/c_k^0))}$$

where the '0' indicates the base year cost. In this case the costs appear as ratios rather than differences but the model is still incremental. For the distribution model the costs are a form of composite cost. The hierarchical structure which is assumed is identical to that in Figure 2 (again, omitting the time period choice). As in the logit model version, trip end constraints are applied to ensure matching of trip attractions with land-use factors.

2.8 Drivers and mechanisms

There are three immediate drivers of change in a transport model. These are changes in:

- Generalised travel costs
- Trip attraction constraints and factors (floor space, employment etc)
- Trip generation constraints (population and car ownership)

If all of these factors do not change from the base then the model will give identical outputs to the base case. When transport policies are applied these will, in general, change the values of the generalised costs of travel. These costs control the modal share between each zone pair and, to some extent, the distribution of trips (which are also constrained or influenced by trip attraction constraints). As the generalised cost of a mode rises or falls ceteris paribus then the mode share will move in the opposite direction. In practice, the costs for all components may change and so the overall effect will depend on how costs change across the modes.

Trip attraction constraints influence changes in the distribution of trip destinations either by simply weighting destinations or by constraining trips to targets values for trip attractions. Combined with generalised costs they determine, within the model, changes in the distribution of trips. The constraints are defined by target values for land-use factors and different factors are related to different purposes. The factors are either exogenously supplied or provided by a land-use model (such as DELTA).

The trip generation constraints require trip productions to correspond to the size, age profile and car ownership distribution of populations within each zone. These may also be exogenously supplied or be provided by a land-use model. The trip generations are calculated by multiplying the populations in particular car ownership and age categories by the appropriate person trip rates (which in STM are assumed to be constant over time). Figure 3 illustrates the various interrelationships between the factors we have just discussed.

So far we have not mentioned that STM is an equilibrium model in the sense that its 'solutions' are supposed give internal consistency between the cost-to-demand relationship (travellers demand curve) and the demand-supply-cost relationships on the supply side of the model. Each run requires the model to iterate to the equilibrium solution (or give up after a large number of iterations). This is illustrated schematically in Figure 3.
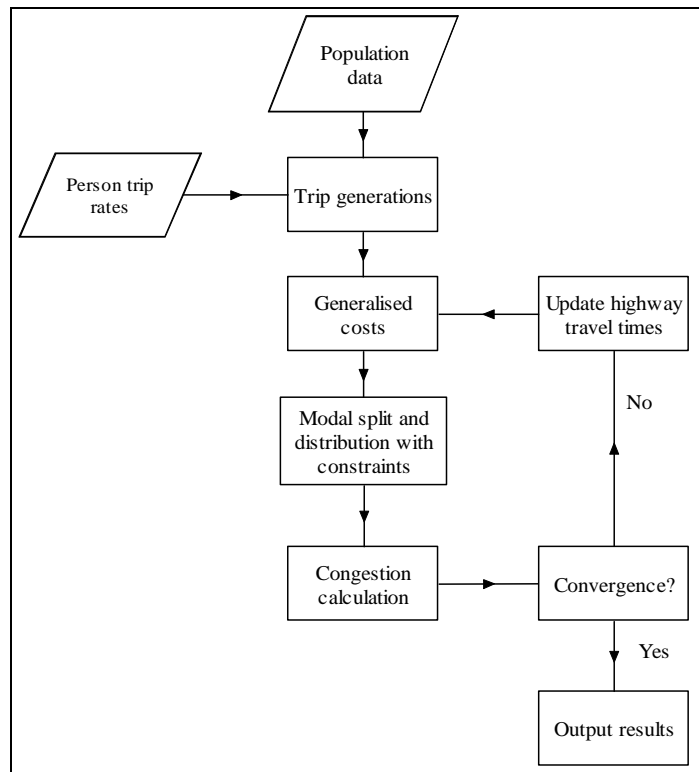

Figure 3. Schematic flowchart of the modelling process

3 The Basic Concept of the Scenario Interpreter

3.1 Interpretation of model outputs

It is normal practice to first run a transport model with policies to be tested under a set of assumptions about demographics and car ownership and then to compare selected outputs (so-called indicators) with those derived from the model for an alternative course of action. In general, the model user seeks to compare all reasonable policy alternatives with one another in terms of indicators which measure the performance of the policies and therefore have a meaningful relationship to the objectives aimed for. It is quite possible that many users of such models do not concern themselves greatly with the reasons for variation in model outputs and more or less mechanically process the results. Two example situations suggest why sometimes they might ask more searching questions. The first case is where planners have certain desired outcomes in mind (perhaps because of a background political or economic agenda) and the model stubbornly refuses to give them the answer they want. The second case is where they are perplexed (not unreasonably) by counter-intuitive results which they fail to explain and judge to be possibly wrong (as they could be). The problem may also be that they have more than one model and inconsistencies have arisen between them.

There are a number of ways in which interpretation can be performed. But first let us consider the word itself. In the context of this report it simply means 'explanation'. Explanations can, of course, can result from a number of factors such as errors in the use the model, data errors and software problems. However, in this report we are concerned only with explanations of outputs in terms of the mechanisms implemented in the model. We next consider the usual methods of interpretation.

The first and perhaps most common approach is semi-intuitive: outputs are interpreted by comparing and contrasting them with those expected from a 'real world' mental model of transport and land-use processes. There are two problems with this. The first problem is the complexity of the processes. The second is that the interpreter reads into the model mechanisms and states which are not in fact implemented within the model.

Another interpretative method (or rather an aid to it) is that of sensitivity analysis. This involves varying an input variable, keeping all others fixed, in order to ascertain the sensitivity of the model to that variable. This can be instructive and help illuminate how the particular variable determines the outputs. It does not, however, necessarily penetrate to the heart of the mechanisms which drive the model and as such is really an extension of the intuitive approach. A further difficulty with sensitivity analysis is the potentially large number of input variables involved and the large amount of computation needed to explore these. A method of prior selection may help here but it does not remove the difficulty.

Yet another approach is one which is not always available to the model user. This involves performing a detailed analysis of what the model is doing when it produced a particular set of results. Such investigations are costly and time consuming and usually require the consultants who built the model to be commissioned to carry out the work which entail require special "diagnostic" reruns of the model and possible software changes.

This brings us naturally to the question of how one might carry out an analysis of the results from a model run, at least semi-automatically, and in a reasonable run time. In short, is a Scenario Interpreter feasible? This is the concern of the next section where the general philosophy of the approach is taken up.

3.2 The feasibility of a Scenario Interpreter

The requirements for a scenario interpreter include:

- Explanations expressed in terms of model mechanisms
- Ease of understanding
- Automatic functioning (to some degree)
- Fast execution

The central issue is how to interrogate the model mechanisms working from model outputs. We have not yet clarified what is meant by explanation in terms of model mechanisms. There are, to be sure, contentious questions about applicability of concepts such as causation and law in social phenomena (transport modelling is of course a branch of economics, a social science). We shall not address these philosophical questions in this report. The STM mechanisms are simply taken to have 'drivers' i.e. variables whose changes generate changes in other variables. Interpretation is then to be understood to be the identification of the drivers of key outputs with some form of ranking attached to them.

A complicating factor in modelling is that the relationships between variables are often circular. A affects B which affects C which affects A and so on. The mechanisms which link these variables in the model do not always have a simple counterpart in the real world (depending on the level of abstraction). Models are often iterative i.e. the cycle of variation (A back to A) is followed as part of a feedback process to solve a mathematical problem rather than as a representation of what happens in the real world.

As an example of this, one might consider the modelling process depicted in Figure 3. This involves estimating demand levels (the mode shares and trip distribution) from initial assumptions about generalised costs (and, in particular, the highway travel times) and then using these demand levels to obtain new highway times (i.e. road speeds). The new times and the old times are compared for 'convergence', and if insufficiently close, new starting values are defined (a weighted combination of old and new) and the process is repeated. This is, in essence, a fairly standard, numerical technique for solving equations and there is no pretence here that actual traveller behaviour is being enacted. Where does all this feature when we consider the problem of interpretation?

We think that the way to deal with this is to focus only on the end of the modelling process. When a model iterates, we should only look at the final solution and not the iterative stages that preceded it. An explanation of outputs in terms of the mathematical techniques used would take us too far from our objective: to explain the outputs according to the model mechanisms in terms which can be understood (and would be of interest) by the model user.

3.3 Example application of the Scenario Interpreter

As an example, let us consider the case where only generalised costs change due to transport policies. The policies which the user can apply to change these costs include:

- Public transport fares, service levels and capacity.
- Cordon charges for private travel
- Parking supply and tariffs.

We leave aside changes in the road and rail network, which can be dealt with using the same general approach. The changes which are endogenously generated (i.e. within the model) are:

- Highway travel times
- Public transport crowding effects.

The purpose of the Scenario Interpreter would be to enable interpretation of changes in the model outputs relative to the base case in terms of changes in the cost factors.

To narrow the example, we may consider as a model output the attractions of trips to a particular zone. These can be analysed into modes and purposes. Let us assume we have established a cordon around the zone and apply a charge whenever a car crosses into the zone or passes through it. We can then examine the changes in attractions to the zone relative to a reference case which we assume here to be the base. We would probably find that the proportion of trips by car had decreased and that public patronage and slow modes had experienced an increase.

We require an answer to the question: to what extent are changes in the attractions are attributable to different cost factors? For each category of travel we need to identify the possible "drivers" and attach to these measures of their significance for the outcomes.

In our example, the answer may not be as straightforward as 'the cordon charge caused car drivers to transfer to public transport' (this interpretation would only be possible with a pure cordon policy anyway). The role of redistribution has also to be considered as well as 'knock on' effects on highway speeds. Conventional sensitivity analysis would have something of value to say but the methodology suggested in this report has additional advantages.

In the present report we will have mainly in mind the interpretation of trip ends, trip starts, zonal speeds and emissions as calculated by STM. The reader may therefore find it helpful to think of these whenever unqualified reference is made to outputs but the methods described here are of more general application.

The next section outlines the method proposed to generate special runs (called bracket runs) which form the basis of the interpretation of model outputs.

3.4 Bracketing methodology

The methodology proposed here for STM is based on the idea that useful inferences about the mechanisms which drive changes in the outputs can be obtained by calculating

the degree to which different factors "contribute" in the final model iteration to those output changes. A useful aspect of the method is the possibility of investigating of the endogenous highway travel times and the crowding effects on public transport. In this report we will focus on highway travel times as this is the most important factor.

The method presupposes that the user has made a standard full run, i.e. the model has run to equilibrium (to the desired degree) with the required policies and planning scenarios and a set of outputs generated. We suppose that the user wishes to investigate the role of a particular factor. To measure the 'contribution' of a factor, it is necessary to rerun the last iteration of model as before but with the change in the factor relative to the base suspended (bracketed). This is called a bracket run and each bracket run is specific to a particular factor. The outputs for the run with and without the factor change could then be compared and the impact (contribution) of the bracketed factor calculated.

The bracket runs will reveal that factors contribute positively, negatively or neutrally to an output. For example, factor x may show a positive effect in the sense that when the change in x is suspended the output decreases a by proportion, say p(x). p(x) provides the measure of the influence of the factor. If other factors are also investigated then a series of values of p are produced.

As an example, if we look at trip attractions T(*,d,-t) (Section 2.5) and examine the factor 'parking costs' we will find that parking costs "account" for a certain proportion p(x) of the change in trip attractions. We may also find that other factors have also played a role (such as highway travel times). Users can obtain some idea of the significance (which could then be supported by conventional sensitivity analysis) of a factor from the weight p(x). This analysis provides aggregate information relating to the trips terminating in a given zone or group of zones. This does not, of course, tell us which movements are the most important for the zone in question. To investigate this it would be necessary to apply the analysis to the actual matrices. This is also possible with the present method but is not pursued in this report.

This methodology is easiest to understand in terms of pure modal share modelling (i.e. leaving out redistribution). If redistribution is included then it would be possible to separate the influence of cost components on the redistribution by bracketing factors in the composite cost used in the distribution model but leaving them intact in the mode split model. Alternatively, the factor change could be suspended in both the modal share model and the distribution model.

In the present report we consider bracketing to suspend changes relative to the base i.e. the reference case is assumed to be the base. This restriction could be lifted and the reference case allowed to be any alternative run (e.g. a do-minimum). In this case bracketing would be relative to this general reference case.

It is important to realise that this is not straightforward sensitivity analysis. In that approach, a factor can be switched on and off to see what effect they have on the equilibrium solution. This has the disadvantage of being time consuming because the model must iterate each time to produce new outputs. The iterative process also obscures the relationship between the outputs and the factor under investigation. In the present proposal only the final iteration is repeated with the factor suspended. If there are several factors then these could be subject to their own separate bracket runs. The outcome of this process will be outputs from a reference case run and from a number of bracket runs. These could then be compared through the STM's Graphical User Interface.

The preceding paragraphs have concerned the factors which form the generalised costs used in the modal split model. It is now necessary to look briefly at the other factors. These are the population factors and trip end factors influencing the distribution of trips.

The population factors are represented by the trip generations. We can devise a bracket run by replacing forecast year trip generations by their base values and then performing the bracket run. The role of trip end constraints could be explored in the same way. Targets for retail floor space and jobs could effectively be replaced by their base values to see what pulling effect they have on travel.

In Section 4 we go on to describe the Scenario Interpreter as sketched above in some detail.

4 The Scenario Interpreter in Detail

4.1 Cost calculations within STM

Crucial to the use of the scenario interpreter is the calculation of costs and the individual cost components in the base and forecast year. Generalised costs within STM are categorised by mode, purpose and time period. For each mode these costs are divided into money costs and time costs. Money costs comprise for car travel:

- Cordon and road pricing charges
- Parking tariffs
- Fuel costs

For public transport modes the money costs are simply the fares charged for a trip. Both car and public transport travel incur the following time costs:

- In-vehicle time
- Access time
- Egress time

In addition, public transport travel incurs a waiting time penalty, which is related to the service frequency. The money costs can be converted into time units by dividing them by the value of time. The converted money costs can then be combined with the time costs to give a generalised cost. This is purpose dependent because the value of time depends on purpose. There can also be an additional penalty which arises from crowding on public transport. For buses, crowding produces a lengthening of the effective waiting time; on rail and underground the penalty is related to the levels of standing in relation to the seating capacity.

Manipulating costs within STM is not as easy as might be first thought. For small STMs it was possible to store in arrays most information and then retrieve it at will. As the demand for larger, more multi-dimensioned models grew it was necessary to restructure the model so that large amounts of information were not permanently stored but instead were used straight away within the model and then discarded. After the model has run through a series of iterations to convergence information required for output (such as matrices) is generated by repeating the final iteration. As information is generated by the model processes it can be printed to file as it becomes available. Some control in the way information is output is achieved by temporary accumulation of data. In addition, when data are not too bulky the traditional use of storage arrays can be applied (e.g. for trip ends). This has limited application – there is no possibility of storing in advance the full forecast cost and trip matrices within STM but the matrices can be written to file.

The way in which STM calculates and stores cost and trip information is described schematically in Figure 1 (Appendix B). This illustrates the main processes involved in calculating generalised costs in the routine NEWTRP – the forecast demand module. This routine is rather complicated so one should concentrate on the features highlighted.

The routine calcgtime calculates the generalised time of travel (weighted combination in-vehicle, access/egress and, for public transport, waiting times) between each origin-destination zone pair in the study area. This is stored in an array gentime(o,d,m,t). The array does not depend on purpose or household car ownership because time costs do not

depend on these factors. Similarly, the routine also calculates the fuel cost for car travel between each zone pair.

Parking costs are estimated at start of model run on the basis of the base case and supply and tariff policies. These are then made available to the next routine we are to consider: setupgcbymode.

The routine setupgcbymode is called twice within NEWTRP: in one section dealing with HB trips and in another treating NHB trips (reduced to a single box in Figure 4). It brings together all the money and time cost components and converts money costs using the value of time. The result of each call is the value of generalised cost for each mode for a given time, origin, destination and purpose. NEWTRP can also accumulate money and time costs (and trips) to be printed out in so-called GCTR files for use by DELTA if it is being used. Within STM the GCTR costs are temporarily stored by destination and mode in arrays which refer to a particular origin, purpose household car ownership and time period. This stored information is thrown away (apart from its presence in a file) as the origin etc cycle. Otherwise there is no other storage of costs within STM.

NEWTRP is called each time the model in the first pass iterates. When the model converges (or reaches the limit for the maximum number of iterations) the last pass is repeated (after resetting variables) to allow outputs to be generated en passant e.g. the GCTR files.

The importance of this preamble is that we can now identify this last step as the key to how the bracket runs are to be generated. Each bracket run corresponds to a call of routines which repeat the final output stage but in each case the values of factors must be set to the appropriate values. Whenever the output stage is run it generates a file foredata.unf. These can be used to extract outputs and compare the results of the reference case with those from the bracket runs.

Figure 4 shows the basic layout of a hypothetical STM with a loop for performing the bracket runs. The number of factors can be read in from a file. An input facility on STM GUI could be used to set the number and define the selection of factors. There are quite a large number of possibilities so it would be prudent not to set too many factors up. Nonetheless this facility will run much faster than conventional sensitivity analysis. The technique can be used for STM runs which cover a range of years. If too much is asked of the system then a large amount of data will be generated and disk space may be put under pressure.

In Section 5 we describe an implementation of the Scenario Interpreter. This differs from the schematic plan in Figure 4 in that the loop of bracket runs is not implemented and each bracket run therefore requires a separate run of STM. This has the disadvantage of the overhead of performing base calculations and base data input in each bracket run but was computationally easier to implement at the pilot stage. The principles of the method here and in Section 5 are however the same.

Figure 4: Schematic Flowchart of model with bracket runs
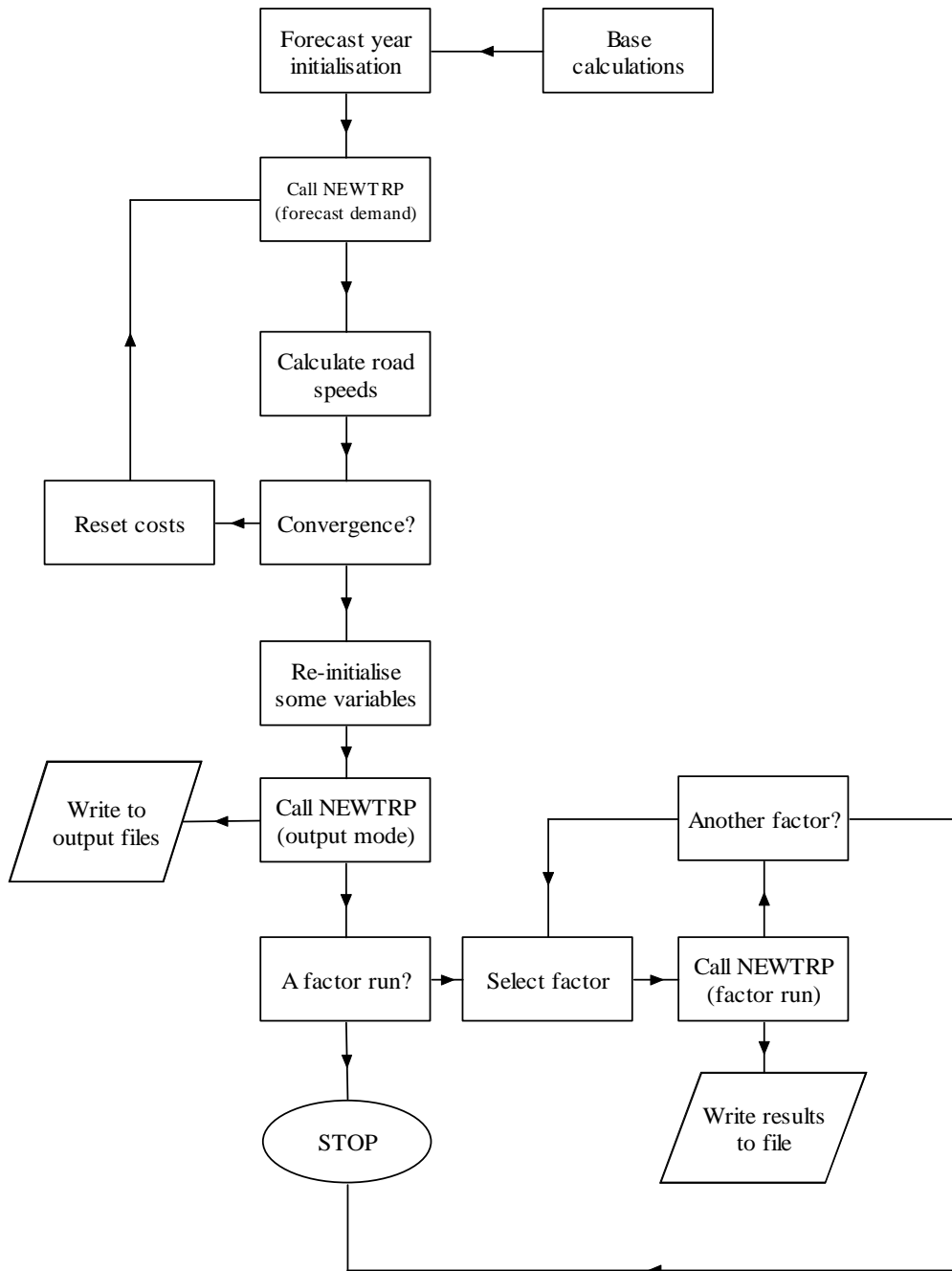
4.2 Basic theory of the bracket runs

4.2.1 Cost factors

By cost factors we usually mean those separate items of cost which make up the total generalised cost for each mode. We can generalise this term to also cover the total cost, money cost and time cost. It is convenient to repeat the equation used in STM Version 4.6 (in slightly recast form) for modal share for the ith mode as given in Section 2.6:

$$p_i = \frac{p_i^0.(c_i / c_i^0)^{-\lambda}}{\sum_k p_k^0.(c_k / c_k^0)^{-\lambda}}$$

The sum is over all available modes. The λ and other variables should have subscripts to distinguish purpose, time and household car ownership. These have been suppressed for the sake of a clearer notation. In a bracket run the forecast cost component for a given mode is replaced by the base value. This gives a new forecast cost which can then be inserted into the new equation. The above equation implies that when a factor for a given mode changes it will affect not only that mode but all other modes. If modes are directly dependent on the same factor then the suspension of the factor change should be applied to all the related modes. This could be considered to apply to bus and car in the model. Bus speeds are assumed to reflect general traffic conditions so it could be argued that highway travel times for car and bus are linked in this way.

The equation for trip redistribution is analogous to the above equation. Bracket runs for this can be performed in analogous way.

It is also possible to consider the role of total generalised costs or money and time cost separately in the same way. This is simpler to program but is less informative than working with individual cost components which have a clear real world meaning (such as travel time). It would be more useful for diagnostic investigations of the model program.

4.2.2 Population factors

The population factors enter the model through the trip generations gen(o,-,t) and the equation:

$$T(o,d,-,t) = gen(o,-,t).pod(o,d,-,t).\text{mod}(o,d,-,t)$$

which was first presented in Section 2.5. The gen(o,-,t) term is determined by the total population in zone o and its distribution across age groups and car ownership categories. The bracket runs can be made by using the base values for the gen() in place of the forecast values. The bracket runs could be applied to specific zones or regions in order to investigate the influence of specific locations. It would also be possible to consider bracket runs related separately to car ownership, age profile and total population (all of which are subsumed in the generations).

4.2.3 Trip end constraints

Within STM, trip redistribution (Appendix A, Section 9) is controlled by cost changes and trip end constraints which require trip attractions by purpose to match purpose-related targets based on changes in land-use factors (e.g. jobs) or where land-use factors act as weights (e.g.. retail floor space). Bracket runs could be carried by setting the values of the weights to those in the base. If constraints are applied (a furness-style technique is used) then it is possible to suspend this by resetting the 'column' factors to unity in order to assess the impact of the land-use changes.

4.3 Accessing the outputs from the Interpreter

There are two main possibilities for extracting data from the interpreter. The first method uses the STM Graphical User Interface (Figure 5).
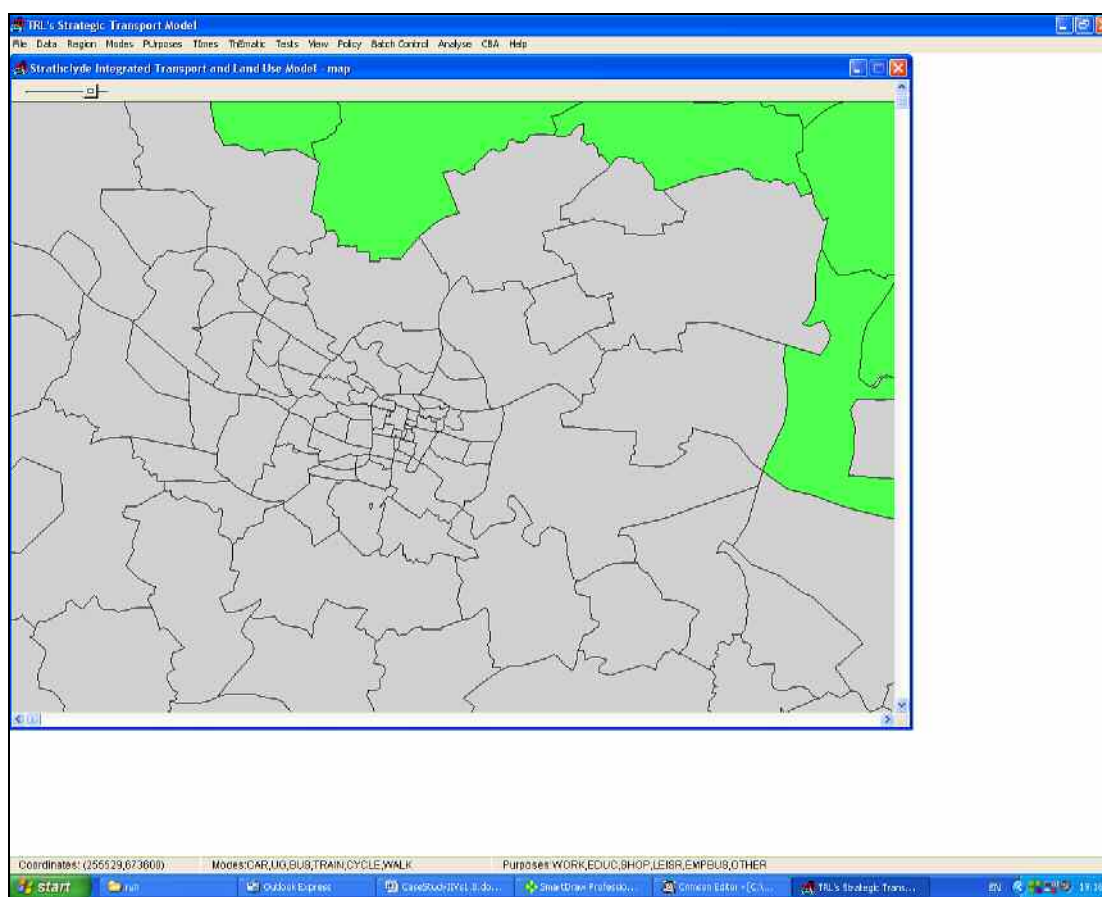
Figure 5: The STM Graphical User Interface

This allows scenarios to be compared with one another (Figure 6). The data used to make the comparison are contained in the foredata.unf files contained in the test folders corresponding to these scenarios. Each test folder contains a subfolder which holds the transport results for the run. There are a number of possibilities for extending the STM GUI to include bracket runs. A neat way of organising the test folders for bracket runs would be to include these as subfolders within the test folder for the corresponding full run. An alternative, and one followed in Section 5, is to treat bracket run test folders as

normal test folders and access these through versions of the standard windows which allow easier identification of bracket runs. As an example, a version of the scenario
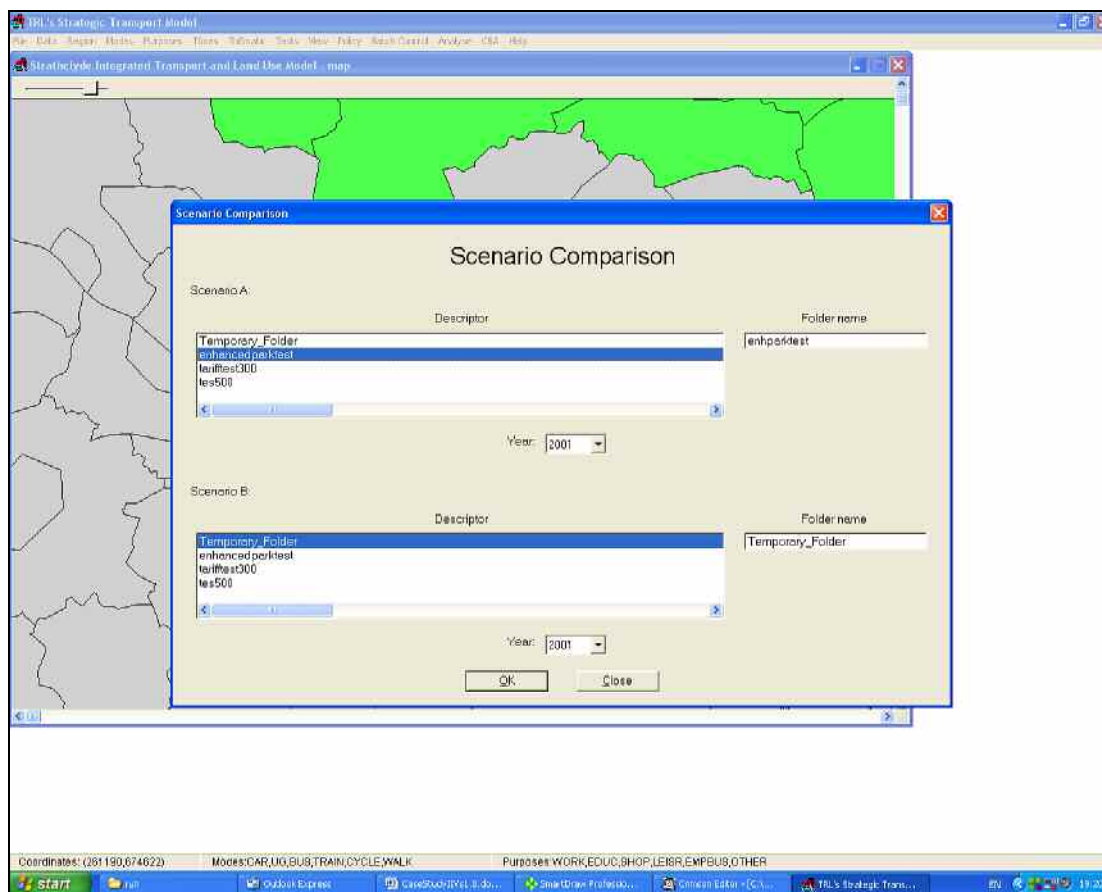
Figure 6: The Scenario Comparison window – this allows selection of scenarios for comparison.

comparison window could be created in which one of the list fields is restricted to bracket runs. The selection of a bracket run would also require selection for a given comparison year. This would allow the standard display features of STM such a bar charts and thematic analysis to be used (see Figure 7).

The other method of analysis is to use the POUTPUTS program. This is an STM utility program which converts to csv format (for use in excel) the STM results in the foredata.unf files in a given test folder. The program combines together the results for selected years and sets these out in selected formats. The csv files are given standard names and are stored in a subfolder 'outsummary' in the test folder. This subfolder is created by POUTPUTS and is not a standard feature of a newly created test folder.
It can be used to generate csv tabulations of the main outputs of the model which then be imported into excel. A simple spreadsheet program could easily be set up to generate tables of absolute and relative differences in the outputs.
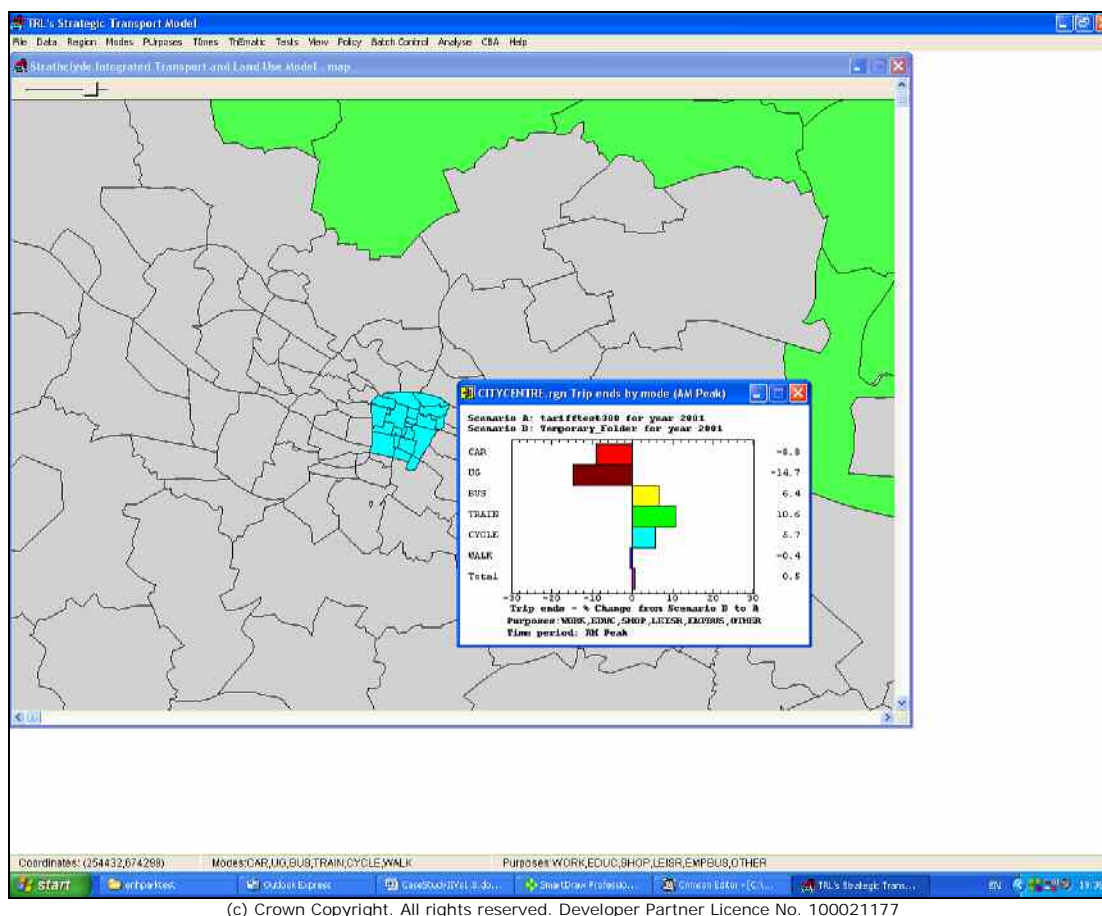
Figure 7 : STM GUI showing the use of bar charts

4.4 Expanding the Interpreter to provide spatial information

4.4.1 Basic problems

The area in which the interpreter, as described so far, is weakest is in its ability to determine the spatial distribution of the main drivers of change. The methodology described so far allows some spatial information to be derived for population factors and trip end factors because we are able to select which zones are to be 'suspended'. This allows some direct aspects of land-use to be explored but not the knock effects which arise through impacts on congestion and travel times.

There are two areas of concern:

- Determining the spatial distribution of the impacts from a particular type of factor (e.g. fares increases).
- Determining the spatial distribution of the drivers of changes in highway travel times.

The information in foredata.unf is not sufficient to allow these questions to be answered. This will require a different technique based on new outputs from the model. As an example of the problem, we may have information from the STM GUI or an analysis based on POUTPUTS on the importance of a factor for changing trip ends, but we do not know how this impact is distributed over the movements to the destination zone. How can we obtain some knowledge of this in a convenient and manageable form?

4.4.2 Spatial distribution of factor impacts

The spatial distribution of factor impacts can only be explored using the full matrices. We have so far described a compact method of investigating factors using only the zone-based outputs (such as trip ends and trip starts); this has the advantage that the bracket runs generate relatively low levels of output data and they can be treated as outputs from conventional runs of STM. If trip matrices are to be output then this should be an optional feature of the interpreter so as to restrict its use to cases where in depth information is required.

The simplest kind of output analysis for trip ends would be a histogram showing the distribution of the contribution of the factor change across all the origin zones. The user could then see which origin zones were the most important contributors to the change in trip ends. A similar approach could be adopted with trip start forecasts; in these cases the distribution would be for destination zones.

This kind of analysis could be carried out for other types of output such as emissions and pcu-kms.

There are clearly many possibilities for programming these features and using computer graphics to highlight the results. In general, it is desirable for the user to be allowed some freedom in the selection of analyses and outputs in order to avoid information overload and to enable the user to employ their own intuition and understanding.

4.4.3 Spatial distribution of congestion effects

The standard outputs of STM (version 4.6) include data on congestion levels (pcu-kms) and road speeds within zones (by road type). Bracket runs for highway travel times would provide information on the impact of changes in road speeds. The zonal congestion and road speed outputs, coupled with highways routes, can then provide some help to determining the causes of changes in the travel times in terms of effects in particular zones making up the route. There then remains the question of what caused these changes. The bracket runs should in turn provide information on these 'route' zones. Clearly there is the possibility of a continuous regression or circularity – this issue needs further consideration.

It is possible that a semi-automatic analysis might be carried out here. At present, we can only sketch such an idea. For trip ends, it might involve:

- Identifying the most important routes in terms of impacts on trip ends in a zone
- Analysing the route to determine the most important routes.
- Using the bracket runs to identify the major drivers of changes on the routes.

We are therefore moving towards an approach in which runs of the basic interpreter generates a database which can then be analysed to provide interpretation of various classes of outputs.

5 An Implementation of the Scenario Interpreter

5.1 Aims and tasks

The theory of the Scenario Interpreter has been set out in the preceding sections. In Figure 4, the various processes have been schematised as if they are contained within one program. These processes include:

1. Running a standard STM run to equilibrium.
2. Running the last iteration of a given standard run.
3. Applying 'bracketing' to variables in last iteration and calculating the resulting model outputs
4. Displaying a comparison of the standard run results with those of a 'bracket' run.

The arrangement in Figure 4 assumes that (1)-(3) are carried out within the STM kernel program. In particular, all the bracket runs are carried together in Figure 4.

In the present pilot implementation we have adopted a less integrated approach but one which has a number of advantages over that illustrated in Figure 4. In the first place, the model uses two versions of the kernel program (which differ only in one module). The first kernel program allows standard runs of STM to be made; at the same time it outputs to the test folder for the each year run a 'restart' file. The second kernel program (kernel2) can read the restart file and, using its contents, restart the last iteration of the standard run. In performing this run it can generate all the outputs produced by the standard run. These are sent to a folder which is automatically generated when the run is set up using the STM GUI program. Bracket runs are performed using the second kernel program. If the second kernel is run without bracketing any variables then it is called the 'reset' run. If bracketing is applied then only one variable can be bracketed in each run of kernel2. In contrast to Figure 4, the different bracket runs are made using separate runs of kernel2. This involves some overheads due the fact that the same base calculations have to be performed each time the kernel2 program is run and the reset file must be read each time. This does, however, have the advantage that the model was easier to set up and test than if a more integrated version had been attempted at the start. In addition, it is a great convenience to separate the standard runs from the restart/bracket runs. Greater efficiency can be achieved in future by allowing the user to choose a selection of bracket runs to be made within a single call of kernel2. This could be implemented by allowing kernel2 to loop over calls of the forecast year module; outputs would then be directed to corresponding folders for each bracket variable.

This accounts for processes (1)-(3). The user does not actually 'see' the kernel programs. Instead the user sets up the runs and examines on-screen outputs via the STM GUI. This allows the user to choose between a standard run and a reset/bracket run, choose policy and planning folders and generate and run a batch program to run STM. The STM GUI will automatically produce a batch program calling the appropriate version of the kernel program. The outputs for standard and bracket runs can be compared using the window facility in Figure 6 or using a special version of this which is explained in Section 5.4.

The implementation therefore entailed the following tasks:

1. Modify the standard kernel program to generate a reset file.
2. Create a 'reset' version of the STM kernel (kernel2 – stage 1).
3. Extend this to include the bracketing of variables as an option (kernel2–stage 2).
4. Develop the STM GUI program to handle bracket runs.

5.2 Kernel program development

This Section is concerned with Tasks 1 and 2. The standard STM kernel (see Appendix A) iterates to an equilibrium or stops iterating after a maximum has been reached. At the end of each pass of the demand models certain new variable values, needed at the start of each pass, are recalculated. Depending on the policies and assumptions about planning data (demography, employment, car ownership) the model will perform a greater or smaller number of iterations (perhaps as many as 50). The demand calculations in each pass are made by a routine NEWTRP (Appendix B). After stopping the iterations, the kernel reruns the final iteration by calling NEWTRP once again but in this call the final model outputs are generated- only outputs related to the state of convergence are generated during the iterations.

To create the second kernel program the variables used to recreate the last iteration have been saved. These allow the costs at the start of the last iteration to be reconstructed without the need to perform all the preceding passes in the standard run. The variables are:

- Zonal speeds by road type.
- Variables relating to crowding on public transport
- Variables relating to levels of demand for parking places

When the standard kernel is run the values of these variables are written to a binary reset file. The second kernel program performs the same base calculations as in standard kernel but the forecast module has been modified to exclude the first pass of the demand model and subsequent iterations and the reset file is read. After some processing of the reset variables, the final call of NEWTRP (generating outputs) is made.

If no bracketing is applied then the values of the 'input' variables to NEWTRP will be the same as in the standard run. We then have made a reset run. The choice of bracketing 'variables' in this implementation is as follows:

- Bus fares
- Subway fares
- Rail fares
- Bus frequencies
- Subway frequencies
- Rail frequencies
- Highway speeds
- Demography.

This list can be easily extended. As stated earlier only one variable can be bracketed at a time. To bracket bus fares simply means to use the base values of bus fares in the bracket run where a bus fares policy has been applied. This removes the direct influence of fares changes but retains other indirect effects such as changes in highway speeds.

If bus fares have not been subject to a policy, so they have their base value, then bracketing will have no effect and one obtains in this case the equivalent of a reset run. Similar comments apply to the other public transport fares and to public transport service frequencies.

'Highway speeds' refers to the modelled speed for cars. Bracketing these simply causes them to reset to the base values. Because bus speeds are a function of highway speeds in STM these also will be recalculated as base values.

'Demography' is in fact the trip generations calculated by the model. In general, zonal populations and car ownership will differ from those in the base year. Bracketing 'demography' simply means using the base year generations in place of the forecast year values.

It is clear that some expansion of the bracketing variables is possible. We could, form example, look at the role of crowding factors on public transport or different elements within parking. In regard to demography we might wish to separate the role of car ownership and population size. The list of variables does not include trip attraction factors – one could, for example, look at them levels of jobs in each zone as a basis for investigation.

5.3 Graphical User Interface development

The value of the scenario interpreter depends in great measure on the ease with which it can be used. In this respect the STM GUI program has an important role to play. There are two functions to the STM GUI:

- Setting up runs
- Viewing outputs

The first modification was in the STM GUI menu bar under the 'batch control' option. The option create batch file menu has been expanded to two options:

- Create standard batch file
- Create bracket batch file

The first option uses the standard features of the GUI (Figure 6). The second option first produces the window shown in Figure 8.
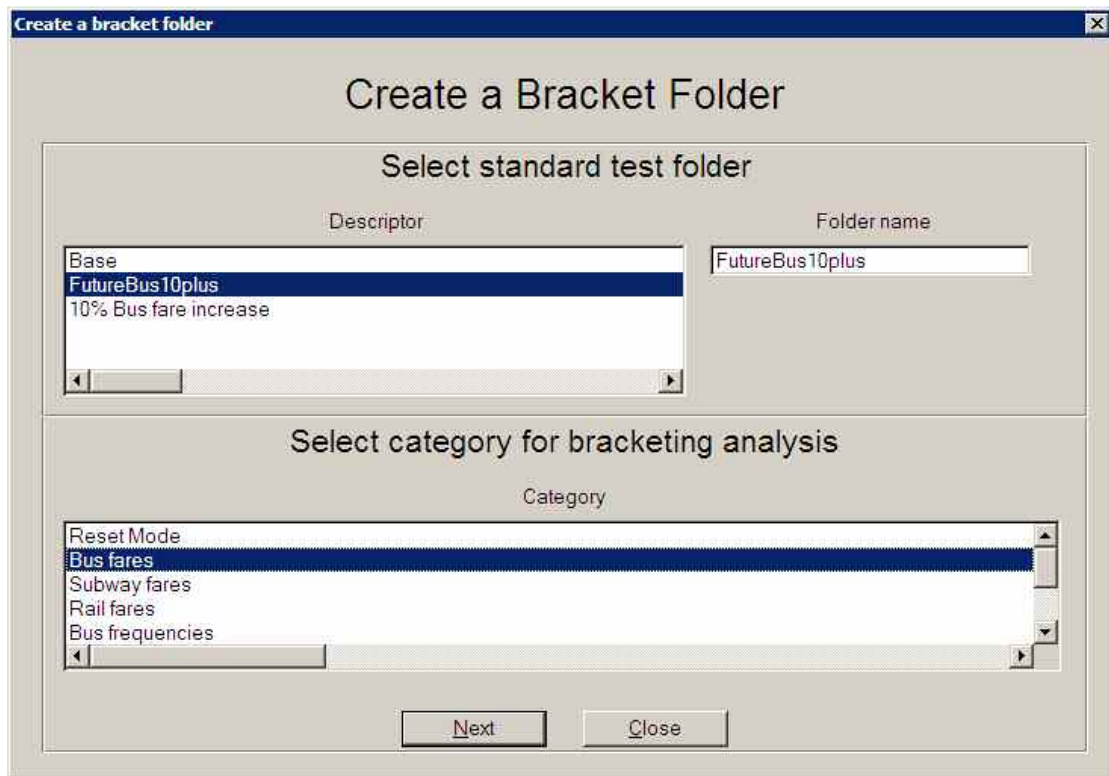
Figure 8: Window for creating a bracket folder

The list box in the upper part of Figure 8 gives a listing of standard runs. The bracket run will be base on standard run selected. In the lower list box are the bracket variable options (including reset). The selections made in these two list boxes determine the name of the test folder into which the results of the reset/bracket run are inserted. The name has the format S_B where S is the name of the standard folder and B is a suffix which identifies the bracket variable (e.g. 'demog 'for demography). The bracket folder has exactly the same structure as a standard folder. If 'next' is clicked then the window in Figure 9 appears. This is a standard STM window which shows for which years runs have been made. Bracket runs can only be made for those cases where a standard run has been made.
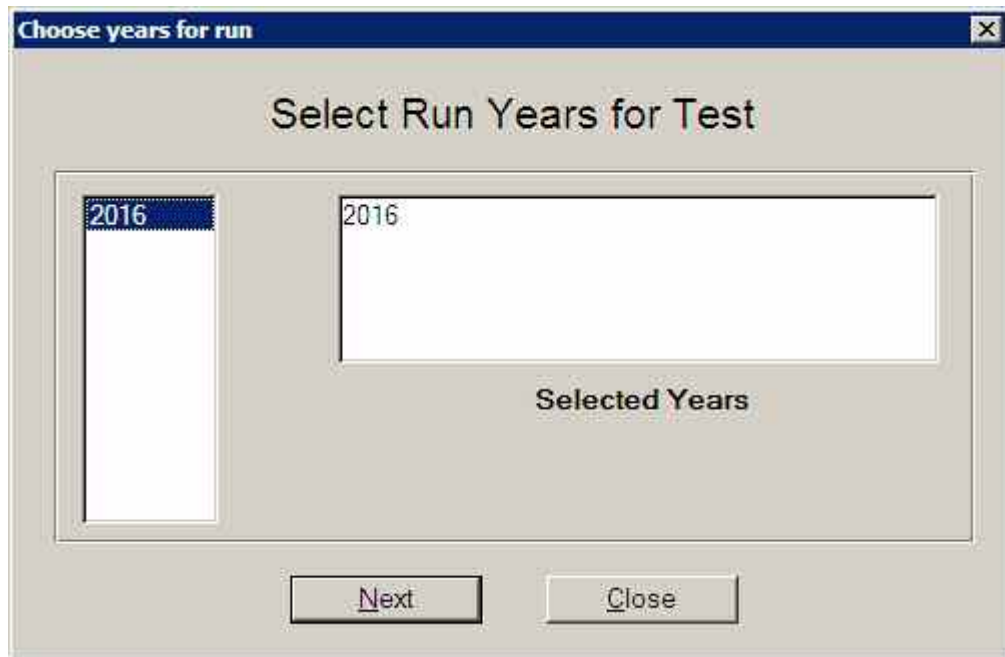
Figure 9: Window for year selection

In the example given only 2016 had been run in the folder 'Futurebus10plus' hence this is the only year which can be selected from the list on the left. Selection of years to run is achieved by clicking on the years required in the left-hand list- they then appear in the box on the right.

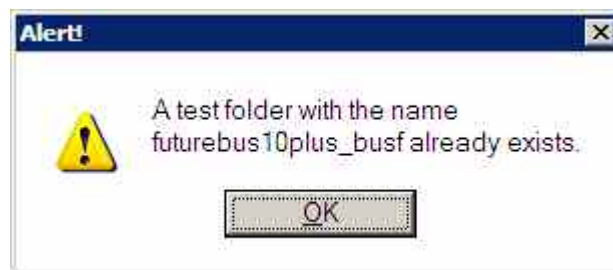If the user clicks next then the window in Figure 10 appears.

Figure 10: Alert message if folder exists.

This indicates that a bracket folder with the same name exists. Clicking on 'ok' produces a window that allows the user to delete the folder (Figure 11). The model is set up so that a bracket run can only be made if there is no existing folder with the same name – no overwriting is permitted as this might result in confusion. If the user clicks 'no' then the set-up procedure for the bracket run is abandoned – it is assumed that the original bracket run is to be retained.

Figure 11. Window for deletion of bracket folder.

Figure 12 shows the result of clicking 'yes' in this case; this window would also appear after the window in Figure 9 if there were no duplication of the folder names.
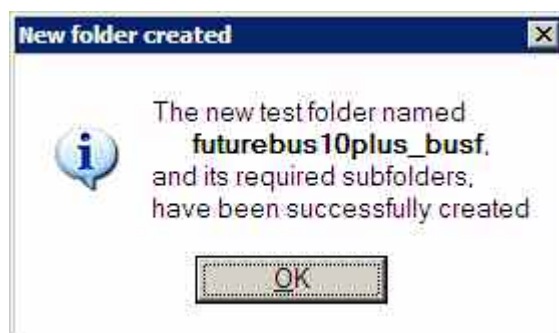

Figure 12. Window to confirm creation of bracket folder
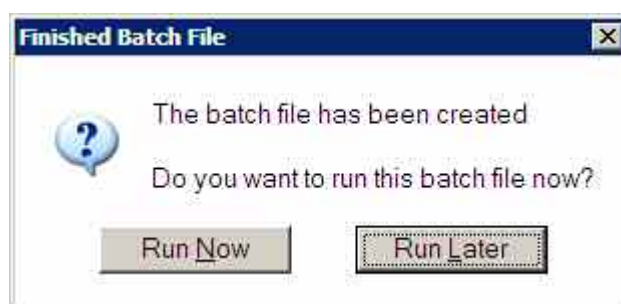
Finally, the user can initiate the bracket run.

Figure 13. Window to run the batch file

The batch file is now set up to use the kernel2 with the policies, networks and planning data of the standard run (in folder futurebus10plus) for the year 2016. The run would also include the same type of bracket run for other years had the standard folder included those years.

The details of the batch file can be accessed using the 'run batch file' option under 'batch control' on the main menu bar.
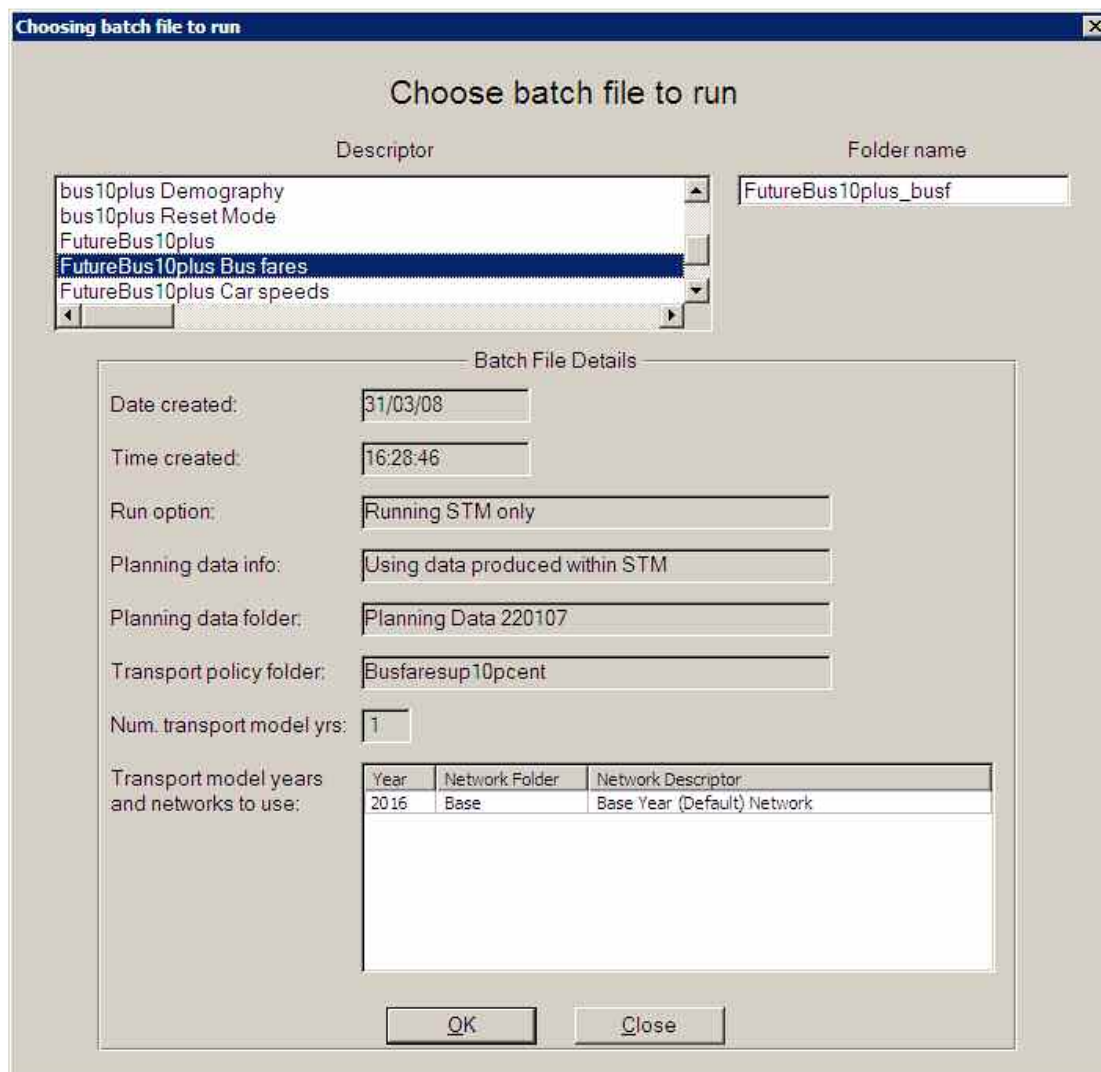
Figure 14.Batch file details for the fares bracket run based on 'futurebus10plus'

The top left list box is used to select a folder on the basis of a 'descriptor'. The actual folder name is given in the top right hand field. It can be seen that the bracket variable is 'bus fares'. The standard run was a global increase of 10% in real terms of bus fares in 2016. The fields below show the folders used for planning data, policies and network. These are taken from the set up for the standard run.

We now need to consider the display of outputs. The outputs for the bracket runs (both in files and on-screen) are of the same type and format as those for a standard run. This means that all the comparisons that are made between standard runs are possible. To assist the choice of folders to be compared a special 'comparison' window has been created. This is shown in Figure 15 and is generated by going to 'tests' on the main menu bar and selecting 'compare bracket results'.
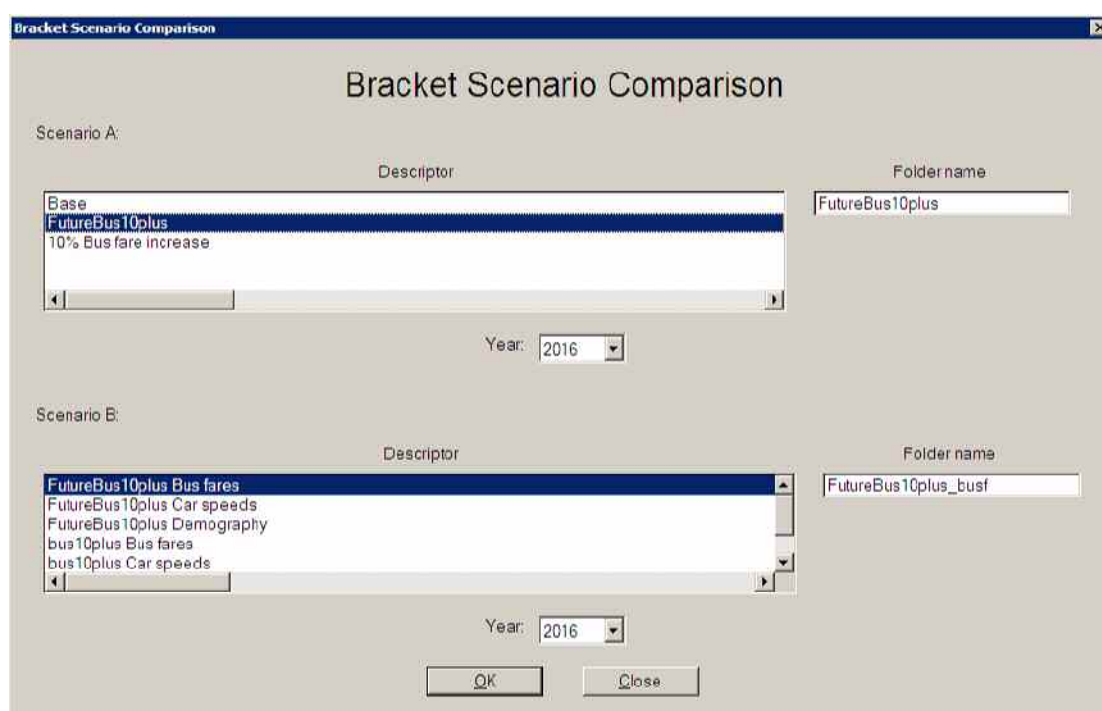
Figure 15. Scenario comparison window for bracket runs

The upper list box is used to select a standard run – only standard runs are listed. In the lower field is a list of bracket run folders. Ideally, the window should be developed further so that only bracket runs corresponding to the selected standard run are shown. The year fields list the years for which runs have been made. The user can also choose 'compare all results' under 'tests'. In this case both fields include a list of all runs. It is therefore possible with this window to make standard comparisons as well as compare bracket runs with one another, if desired.

5.4 Example outputs

The following results are outputs from the DISTILLATE version of STM (version Dist 2-4.6). This was developed from the version used in Case Study I (trip chaining) which in turn was developed for DISTILATE from version 4.6 of STM, which has been superseded by later versions of STM (Version 5.2 is the latest, at the time of writing this report). Results from these DISTILLATE versions should not be taken to be forecasts by Strathclyde Partnership for Transport or represent any policy of theirs or any other body – this is particularly because the versions used in DISTILLATE were specially constructed as experimental platforms for model development and therefore our primary concern was not the accuracy of the base matrices and the calibration of the model. The results presented are based on the prototype and therefore should be regarded as provisional. Development of the model is likely to continue.

EXAMPLE 1: 10% Fares increase – 2001

In this example we use the planning data for 2001 (the base year) but run the model with a global 10% increase in real fares. We wish to investigate the changes between the base (the base run) and the run with a 10% fares increase (the policy run). Figures 16-19 show the impact for the 'Glasgow Conurbation' (in blue) for trip ends by mode for various run comparisons. Figure 16 is an illustrative screen dump showing the Glasgow

Conurbation map and the bar chart giving relative changes between the policy and base runs. The bar chart is reproduced on a larger scale in Figure 16a. It can be seen from this that the 10% increase in fare has reduced bus usage by about 4.7% corresponding to an elasticity of -0.47. The usage of other modes has increased, with car use increasing by about 0.7%. Figure 16b shows the corresponding absolute changes between the policy and base runs produced by the bus fares increase.
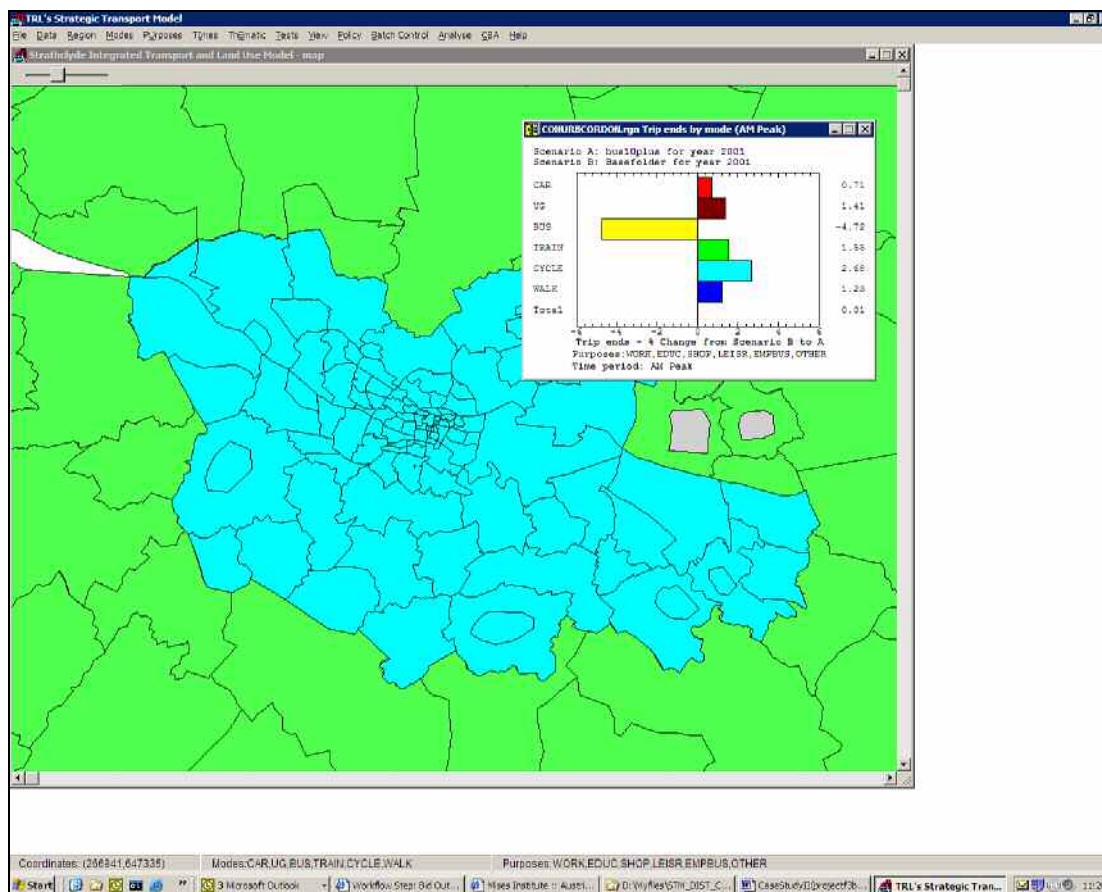
Figure 16: Screen dump showing Glasgow Conurbation and comparison between base and policy run
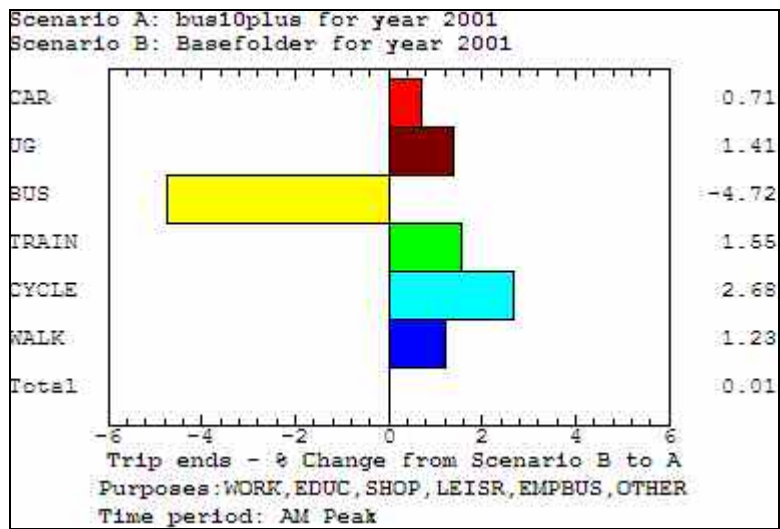
Figure 17a: Relative change in mode usage resulting
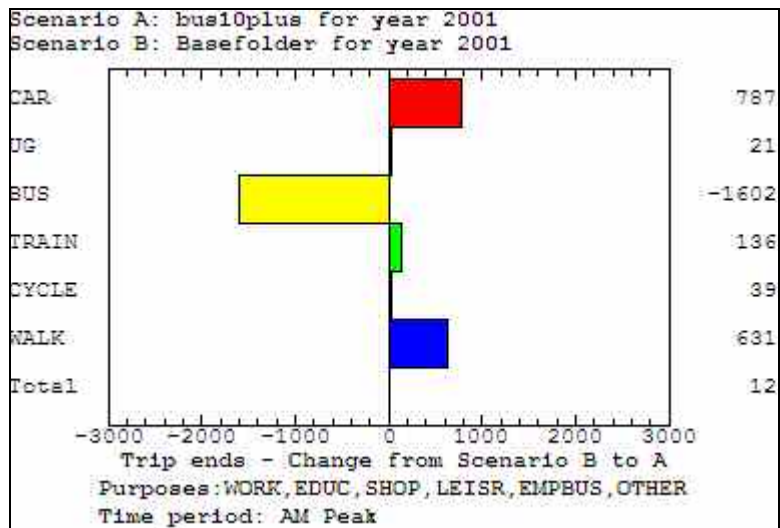from a 10% fare increase

Figure 17b: Absolute change in mode usage (person trips per hour)
resulting from a 10% fare increase

Figures 18 and 19 compare results for standard policy and base runs with those from bracket runs for bus fares and highway speeds. The effect of bracketing out the bus fares is shown in Figure 18. 'Bracketing' bus fares means that the money cost of bus travel will be equal to that in the base. In the present setup, the travel times will have their final forecast values. Figure 18a is shows the relative difference between the policy run (made with a 10% increase in bus fares) and the same run with bus fares bracketed. The policy run has 1956 bus passenger trips/hr more than the bracket run but 973/hr more person car trips. The only difference in the inputs to the two runs is the impact of speed changes for both car and bus travel (which, in the model, depends on the prevailing car speed). In Figure 17b a loss of 1602 trips/hr by bus was generated. This is reflected in Figure 18b which compares
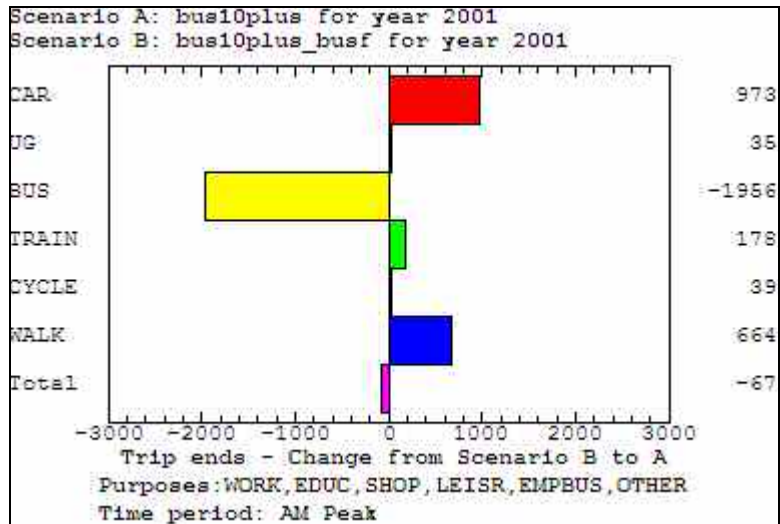
Figure 18a: Absolute comparison for policy bracket run of bus fares
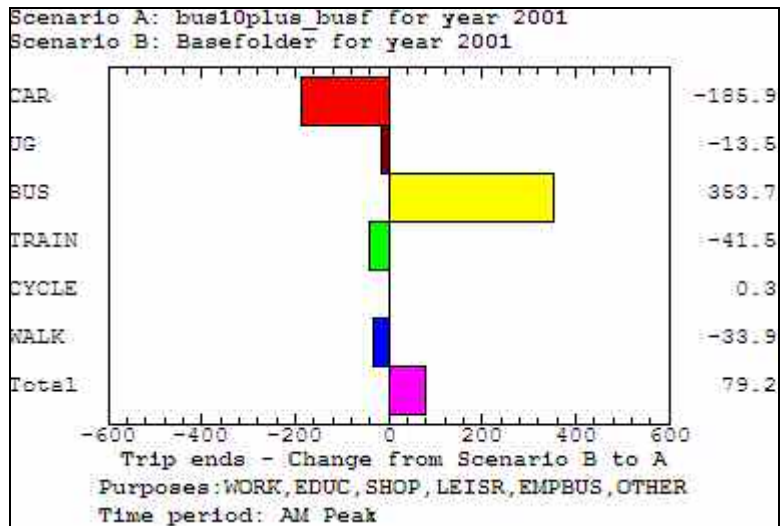with a standard run with 10% increase in bus fares

Figure 18b: Absolute comparison for policy bracket run of bus fares
with the base run

the policy bracket run for fares with the standard base. This shows that the speed impacts left over from the standard run produce a definite shift from car to bus. The corresponding relative changes in trips are shown in Figure 18c.   Note that the change in bus trips generated by the impact of car speeds (Figure 18b), 354 trips/hr, can be subtracted from the bus trips in Figure 18a to give 1602 trips/hr , which is the difference between bus trips between the policy run and base run. We can thus see the role played by speed changes separately from the direct impact of bus fares.
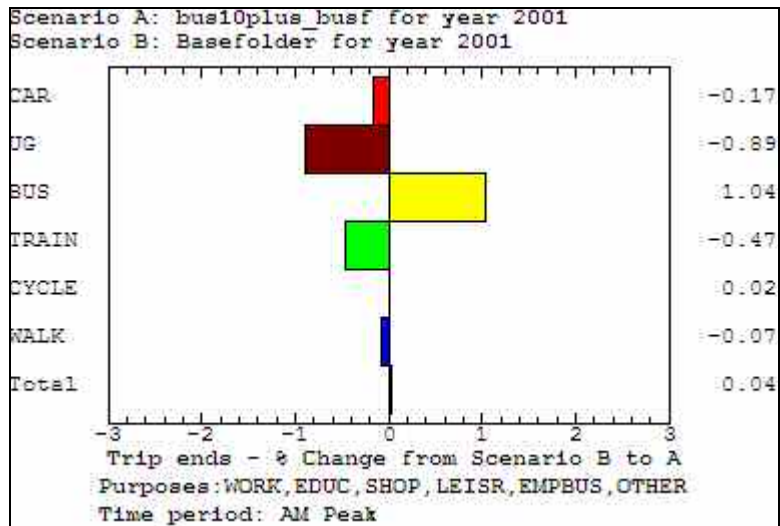
Figure 18b: Relative comparison for policy bracket run of bus fares
with the standard base run
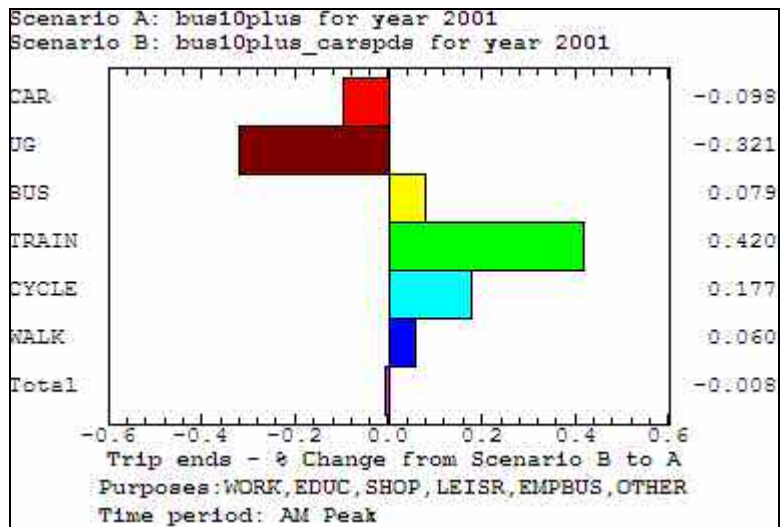
Figure 19a: Relative comparison for policy bracket run of highway speeds
with a standard policy run with 10% increase in bus fares
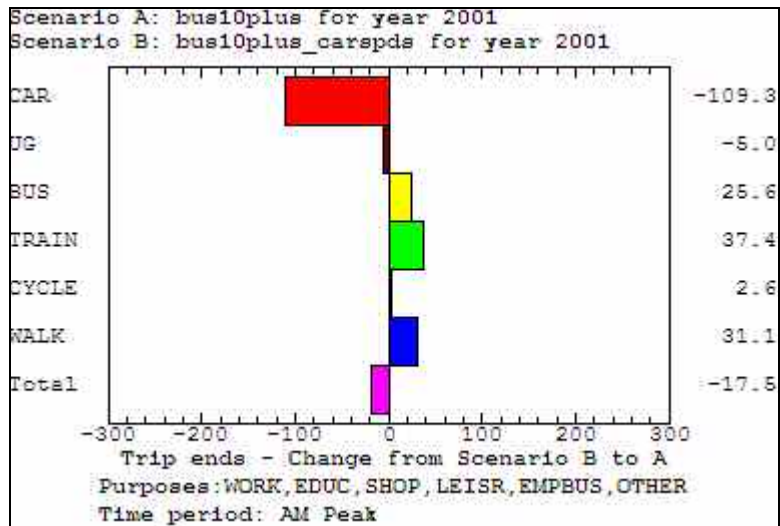
Figure 19b: Absolute comparison for bracket run of highway speeds
with a standard run with 10% increase in bus fares

EXAMPLE 2: Change in planning data set.

In this case we compare a 2016 run and a 2001 base run for the Glasgow Conurbation without applying any transport policies. The impacts therefore arise from the changes in land use. The 2016 planning set was based on multiplying base planning factors by TEMPRO-based growth factors.
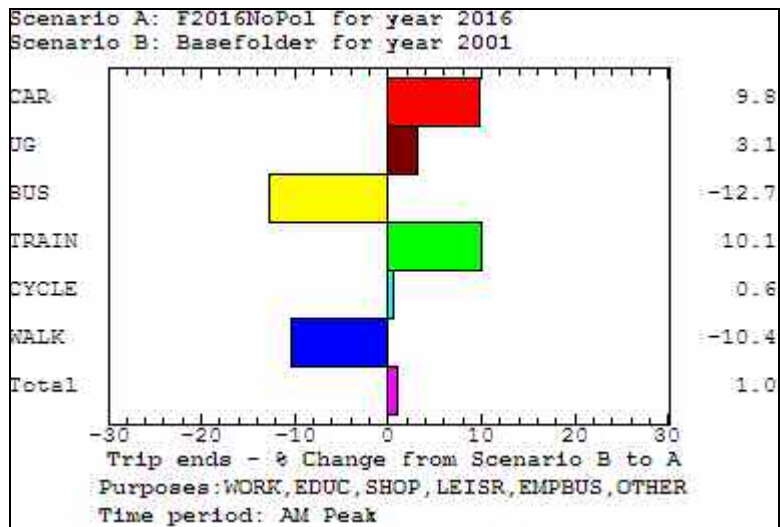
Figure 20a: Absolute comparison of the 2016 run with the 2001 base run
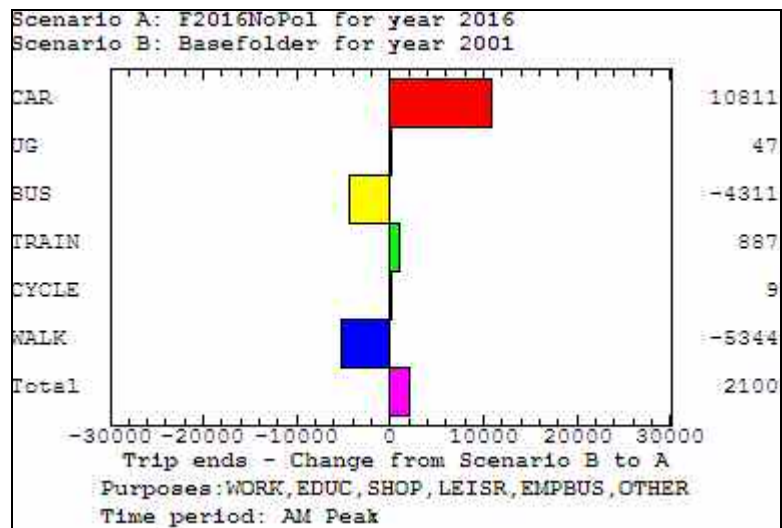
Figure 20b: Relative comparison of the 2016 run with the 2001 base run
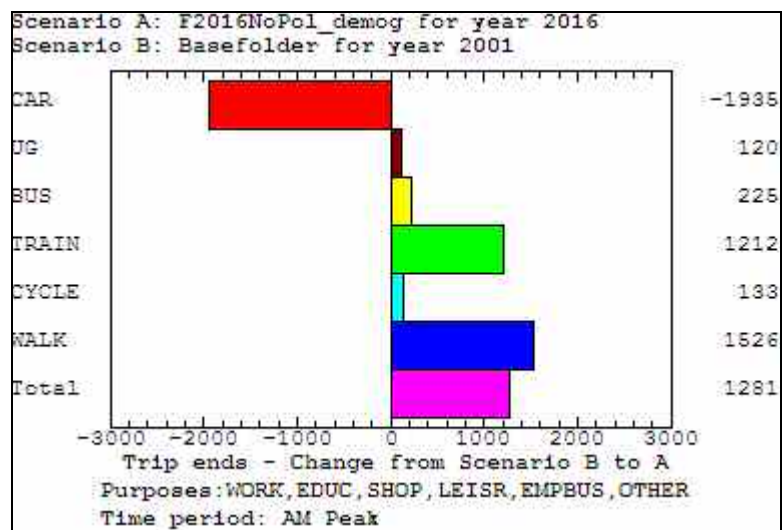
Figure 21a: Absolute comparison of the 2016 demography bracket run
and the 2001 base
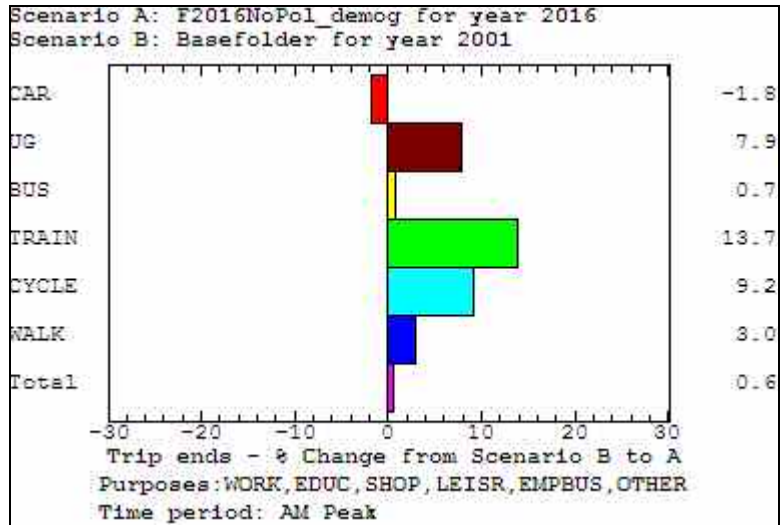
Figure 21b: Relative comparison of the 2016 demography bracket run
and the 2001 base

Figure 21 compares the 'demography' bracket run in 2016 with the base run. This shows the impact of the speed changes and crowding effects on their own. It would appear that speed reductions have driven down car trips by about 1.8% and have displaced trips to public transport (particularly rail) and walk. Transfer to bus is relatively weak, perhaps due to crowding effects.
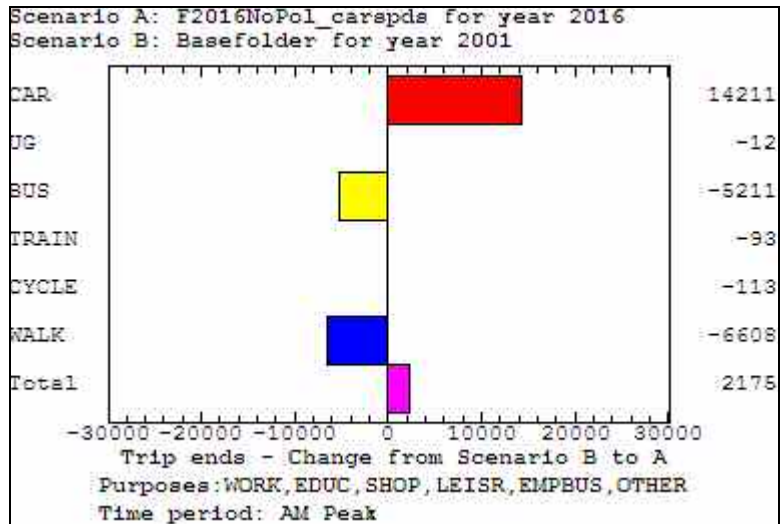
Figure 22a: Absolute comparison of the 2016 highway speeds bracket run
and the 2001 base
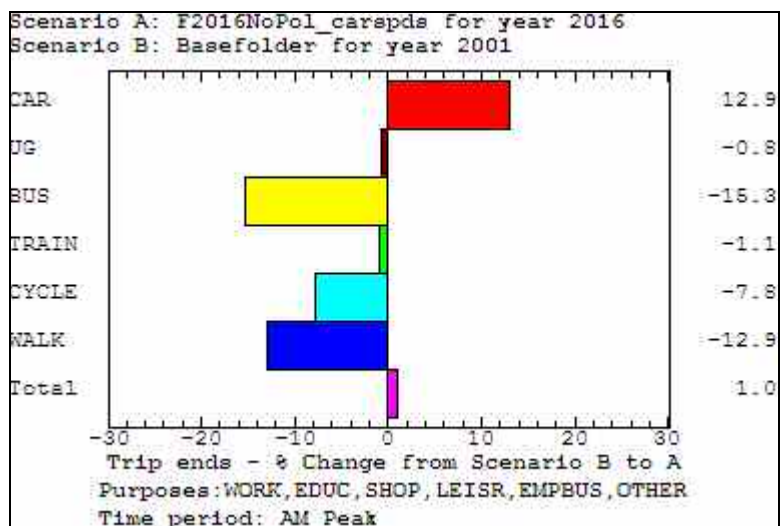
Figure 22a: Relative comparison of the 2016 highway speeds bracket run and the 2001 base

Figure 22 shows the corresponding results for the bracket runs for highway speeds. In this case the changes in demography have been retained but the highway speed changes suspended. The result is a large positive change impact on car in contrast to the negative impact when demographic changes were bracketed. The changes in person car trips and in bus passenger trips in Figures 21 and 22 do not sum to give the changes observed in the standard run (Figure 20). For car, the sum is a gain of 12276 trips per hour in the am peak (13% higher than in Figure 20); for bus, a decrease of 4986 trips per hour (about 6% less than in Figure 20) is obtained. This case has not been analysed in detail (it requires the bracket variable list to be extended) but it seems that the gains observed in this run in rail and underground do not come from the changes in 'demography'. This could be explained by supposing that the principal changes occurring under 'demography' are in car ownership and that rail use is rather insensitive to the change in car ownership mix, whereas bus patronage is rather more sensitive (decreasing as the household car-owner ship mix changes). As indicated, these issues could be resolved by extending the list of bracket variables to provide more diagnostic information to work on.

6 Future development possibilities for the concept

Several areas of work merit further consideration:

- The bracket variable list needs to be extended to include the public transport crowding factors and the total travel times for public transport.

- In addition, the 'demography' options needed to be extended to allow bracketing of household car ownership, total population and population composition.

- The model needs to include bracketing for trip attraction factors such as jobs and retail floor space and the cost-based distribution mechanisms in STM also need to be included within the interpreter's framework.

- The selection of bracket runs needs to be made more flexible so that sets of variables can be bracketed together rather than individually.

- The Interpreter needs to be developed further along the lines of an analyser + database, with algorithms to perform searches to identify the most important drivers. The current proposal is limited to a toolkit of methods which require the user to do much of the work.

- The work covered in this report is capable of considerable extension into the realms which touch on the fields of expert systems and artificial intelligence. At the very least it would be desirable to develop an interface capable of translating analyses of outputs into natural language.

7 Summary and conclusions

This report has described a concept for a Scenario Interpreter and a simple prototype has been developed and demonstrated. Its purpose is to provide the user with an evidential basis for the interpretation of scenarios generated by the TRL Strategic Transport Model (STM). This will allow users to identify the likely drivers and mechanisms responsible for model outputs under particular land-use and transport policy assumptions.

We have identified the key techniques to be employed in the interpreter; central to the method is a technique of using a first order sensitivity analysis applied in the final iterative pass of the STM (as distinct from one based on simple re-running of the all model iterations). This has the advantage of reduced run times and a clearer relationship between outputs and model mechanism.

The report points to possibilities for further development of the software implementation of the model and we conclude that the technique is capable of considerable elaboration with automatic algorithms to carry out the analyses. We also look forward to a version of the interpreter which can provide results of analyses in natural language.

8 References

Aramu A, Ash A, Dunlop, J and Simmonds, DC, 2006: SITLUM : The Strathclyde Integrated Transport/Land Use Model, Proceedings of the EWGT 2006 Joint Conferences, Bari, Italy, pp503-510

Appendix A : TRL's Strategic Transport Model

A.1 STM Model version

The platform for this study was TRL's Strategic Transport Model (STM) for Strathclyde . This has been used by SPT on their PCs for a number of years. The version adopted (Version 4.6) as a starting point uses fixed zonal routes between origin and destination zones and highway speeds are estimated within each zone using road-type dependent speed-flow equations. This version was also developed to operate in conjunction with the DELTA land-use model created by the David Simmonds Consultancy (www.davidsimmonds.com) so as to form a Land-use and Transport Interactive (LUTI) system . The LUTI model developed for SPT is called the Strathclyde Integrated Land-Use Model (SITLUM) (Aramu et al, 2006). We do not use DELTA with STM in the present study as its use does not add anything to the development or assessment of the trip chain model. In its stand-alone mode STM forecasts travel patterns using exogenous (user supplied) planning data.

Since Version 4.6, STM has undergone considerable development on behalf of SPT. The new model can now model, within STM, congestion on individual road links using updatable routes and routing factors imported from a SATURN model using a bufferised highway network. The demand models can also handle three time periods (a pm peak has been added) and a choice of hierarchical logit structures for demand modelling is available. The new SITLUM arising of this STM development can provide planning data and growth factors to a large assignment model for detailed network forecasts. The modelling techniques described in this report can also be applied to this new version of STM.

A.2 The STM programs

The STM program system comprises two programs:

- A transport model proper (the kernel STM program) which forecasts in a given forecast year under prescribed policy and demographic conditions.
- A graphical user interface (STM GUI) which allows the user to set up policies and planning data sets, run the STM kernel and view output graphics.

In general, 'STM' will refer to the STM kernel. In Section 5.3 of the main body of the report the STM GUI is discussed briefly. The following sections concentrate on the kernel program.

A.3 Zone system and attendant assumptions

STM is concerned with modelling trips between zone pairs. Each zone ideally represents an area of uniformly distributed land uses (preferably a single land use) and trips are treated as having their ends uniformly distributed within the origin and destination zones. STM also makes certain "averaging" assumptions about congestion levels within each zone. The model designer must choose a zone system (i.e. their number, shape and size) which limits the computational effort (which is proportional to the square of the number of zones) but allows a sufficiently high spatial resolution.

Typically, zones are wards or convenient subdivisions of these; sometimes the zones may have been based on the zones of another model in order to achieve conformity with it.

The zone system comprises internal zones and external zones. As these terms suggest, internal zones form the core of study area. External zones are located at the periphery and serve to represent the influence of areas at the edge of or at some distance from the core area Movements between internal zones are fully modelled by STM i.e. all the trips between internal zones and all other zones are fully accounted for by STM. Trips between internal zones and external zones are modelled but STM does not model trips between external zones.

A.4 Trip representation

STM works in terms of person trips made per hour averaged over a suitable time period. Vehicle flows and their implications for congestion are obtained by division of person flows by vehicle occupancies. Freight demand is not modelled but a congestion effect due to freight can be added to that due to cars and buses by estimating the pcu-km contribution as a fraction of the total.

Trips are segregated by

- Mode
- Purpose
- Household car ownership
- Time of day

Modes

The mode used is the "main" mode for the trip. In practice trips will consist of a number of stages each with its own mode. Thus a traveller may walk to bus stop, catch a bus, and then walk to their final destination. In this case the trip would have as its main mode "bus" and the walk trips before and after would be the access and egress elements for this trip. Clearly more complex trips are possible in which a clear main mode is not possible. In this case the bus trip above might be followed by a rail trip of similar length. STM currently does not incorporate such trips explicitly. All trips are reduced to the "main mode" type.

The standard modes in STM are:

- car
- bus
- rail
- underground
- walk
- cycle.

Taxi and motorcycle can also be included.

Purposes

Some trips may have stages with different purposes: e.g. dropping off children at school on the way to work. These are not explicitly modelled in STM. Each trip is idealised as having a single main purpose (the activity at the destination).

The home-based purposes included are

- work (mainly commuting)
- education (mainly school trips)
- social/leisure, shopping
- employer's business
- other

STM models two non-home-based trip purposes:

- employer's business
- other.

Household car ownership

Travellers' response to travel conditions will be influenced by a number of factors related to their socioeconomic status. Household car ownership serves as a measure of income and car availability. In the current STM we have used household car ownership of travellers i.e. the number of cars possessed by the household to which the traveller belongs. This currently has three levels: 0, 1, and 2+.

Time periods

Version 4.6 of STM currently has two time periods:

- am peak period
- inter-peak period

The am peak can be an average of the peak hour (0800-0900) where traffic levels tend to be highest or an average over 0700-1000. The inter-peak period is modelled as the average flow in the period 1000-1600. In the SITLUM STM, the am peak period is 0700-1000.

The home–based trips we model in the STM demand models are the "outward" stages i.e. the trip from home to the destination. Return trips can be estimated by applying trip-purpose-dependent "return factors" to the outward flows. These give the proportions of return trips in a time period generated by outward trips in the same or an earlier period.

Version 4.6 of STM does not model an evening peak (e.g. 1600-1900) (this is done in later versions); most trips at this time are probably return commute trips. STM could be extended to include this but some refinement in the model would be necessary if evening peak specific policies were to be applied.

STM models an average weekday (Monday-Friday). We have yet to construct a weekend travel model. At weekends the purpose mix and time distribution of trips will be very different from an average weekday.

A.5 Demographics, car ownership and employment

When DELTA and STM are run together the planning data and employment levels are provided to STM by the DELTA land-use model, both for the base and forecast years. DELTA incorporates its own sophisticated car ownership model which uses larger numbers of traveller/household categories as required by the land-use calculations.

A.6 Trip matrices

Within the STM, the base trip matrices are represented by

1. zonal person trip rates ("trate" matrices)
2. base planning data (exogenous or from DELTA).
3. modal split matrices (split matrices)
4. distribution matrices (pod matrices)

The 'split' matrices give the modal shares for each zone pair, purpose, household car ownership category and time period. The 'pod' matrices give the distribution of trips over destination zones by purpose and time period for each zone considered as an origin. Home based matrices can be generated by calculating trip generations from 1 and 2 and then applying split factors (3) and distribution factors (4). Non-home-based generations are then derived from home-based attractions and trip rates in (1) and the matrices finally obtained by applying (3) and (4) for non-home-based purposes.

The forecasting mechanism uses forecast year demographic and car ownership data, employment and travel to work data to drive the new trip generations and attractions using the base person trip rates; forecast travel costs are used to estimate new "split" matrices from the base "split" matrices in the modal split model and new "pod" matrices from the base "pod" matrices in the redistribution model.

A.7 Mechanisms - summary

The basic function of STM is to predict forecast trip matrices under policy and planning scenario assumptions. From the predicted travel costs measures of congestion and accessibility are possible. From speed and flow information predictions can also be made of vehicle emissions levels and "road safety".

STM is incremental in the sense that forecast trip matrices are calculated starting from base trip matrices, which must be supplied in input data files, and base cost matrices, which are constructed from input cost data files and forecast year policies and planning data.

STM is also an equilibrium model in the sense that the basic processes (excluding home-based trip generation- see below) are iterated until the starting costs (prior to modal split/redistribution) and output costs are sufficiently coincident.

STM incorporates four basic elements to make forecasts.

- Trip generation

- Modal split
- Redistribution
- Congestion calculations

Home-based (HB) trip generation is based on demographic and car ownership data and trip rates per person (in different household car ownership groups). These are created at the start of the modelling process. Non-home-based (NHB) trip generations are calculated from home-base trip attractions (which serve as trip generators for NHB trips) to each zone and trip rates per generator trip. This is carried out in the Modal split/Redistribution

The modal split/redistribution is performed "simultaneously" i.e. they occur within each iterative pass. After this each modal split/redistribution the STM calculates the traffic levels (pcu-kms in each zone). These are then used to calculate highway speeds which can then be compared, as a convergence test, with the speeds used at the start of the iteration.

A.8 Trip generation

Home-based trip generation is calculated within for each zone within the study area. This is the total number of HB trips starting within a zone by purpose, household car ownership and time period. The usual approach is to derive, for each zone, person trip rates, i.e. trips per person per hour by purpose, household car ownership and time period from the base year trip matrices and base year demographic data. These rates are calculated outside STM and are provided by the model builder in an input file. The normal assumption (in the absence of other data) is that the person trip rates (disaggregated by purpose etc - not the total trip rates) are stable over time. This is an assumption one might wish to modify. Changes in trip making (generation) occur because population changes and car ownership changes, shifting people from one category to another. The disaggregated trip rates could be made time dependent e.g. by applying time trend factors (to represent social trends) or by allowing travel costs to modify the rates (thus making HB trip generation part of the iterative loop).

In obtaining these trip rates, the total base year generations by purpose, household car ownership and time period are calculated from the base trip matrices and then divided by a purpose-dependent "generator" population to give the trip rates per person. The "generator" populations are chosen to correspond to the main generators for the particular trip purpose e.g. working age adults for work trips.

Non-home-based trip generations are treated in an analogous way to the HB trip generations except that in this case we do not have a generator population of people resident in a zone. Instead the "generators" are work and total attractions to each zone. As in the case of HB trips, a base calculation is made outside STM using the base matrices (for NHB trips). These provide base values for the generator attractions and are used to calculate trip rates per generator, analogous to person trip rates. These are supplied to STM in an input data along with the HB trip rates, discussed earlier. In the forecast case, the forecast year "generators" are calculated within each iterative modal split/redistribution pass. These can then be applied to the NHB trip rates to obtain the NHB generations. The user might consider alternative definitions for the NHB trip rates and the generators. Other NHB purposes might be considered.

A. 9 Modal split and redistribution

These processes fall within the iterative loop of the model. The calculations are based on changes in generalised cost relative to the base. Generalised cost in STM is a combination of time and money costs in which time cost have been monetarised using value of time (VOT). The VOT are for the UK and were taken from figures provided by the UK Department for Transport. The elements in the generalised costs include access and egress costs and travel time. Private travel by car includes fuel cost – using formulae for fuel consumption given by the UK Department for Transport.

Car travel also includes possible cordon charges and parking charges. Parking charges are a combination of the money charge and costs relating to searching and accessing a parking place. Public transport costs do not include fuel cost or parking costs but there are elements for waiting time and fares. The model builder provides data on fares in the base – policy factors can then be applied to obtain forecast fares. The generalised cost for bus allows for bus crowding (a modification of bus frequency) and rail crowding (related to standing).

Also, it should be borne in mind, as indicated above, that the demand response is based on what is effectively the outward trip. It would also be possible to base the response on an average of the out and return trip.

The modal split model redistributes total trips between each zone pair over the different modes in the base in accordance with the way in which generalised costs have changed from the base. The approach is incremental in the sense that it automatically reproduces the base market shares. Calibration is achieved by running the model and performing sensitivity tests for fuel price and public transport fares. The outputs demand responses for the study area as a whole are then compared with published elasticities (UK-based values). Parameters (set in a data file) can then be adjusted to produce a match with the observed elasticities. Some comments are necessary here. The current STM practice relies on the model structure to imply most of the elasticities for different costs elements – comparatively little data exists for some of these. The spatial variation of elasticities e.g. with distance is also implied. The parameters used in the modal split need to be disaggregated by purpose, household car ownership and time period – in practice we have usually worked from a full starting set of parameter values, taken from earlier TRL studies with STM and adjusted these with factors independent of purpose and car ownership so as to obtain a full set with the calibration of each new model.

Redistribution over trip destinations is based on changes in generalised costs from the base and changes in trip end factors. The impact of cost changes is modelled using "composite costs" (calculated over all modes) between zones in a model similar to that used for modal split except that "mode" is now replaced by "destination zone". Changes in costs therefore cause the allocation of trips to destination zones to change. The sensitivity of redistribution to changes in cost is controlled by a parameter which is currently set for work and non-work trips. The distribution of trips is also controlled by "trip end" attractions. Currently this control is only applied to work trips. In this case, changes in the distribution of jobs require the distribution of work trip ends to be distributed in a corresponding way.

A.10 Highway and Rail Routing

Highway

STM does not itself perform an assignment calculation for highway movements by car or bus. Normally, a calculation is made outside of STM of the likely strategic routes between each zone pair using a simple assignment. Each route is converted into a zonal route i.e. the sequence of zones through which the route passes. For each zone, the predominant road type on the route is also recorded along with the length of road through the zone. The model works in terms of motorways, a roads and other (essentially local roads). These are distinguished according to whether a road section is located in an urban or rural setting. These routes are used in STM as the basis for estimating the travel times between zones and it calculating the spatial distribution of congestion.

This information is stored in an STM "network" file. Often these routes are retained through all forecasting years for the purposes of estimating road speeds and journey times. It is also possible for STM to allow the network to be changed in forecast years so that a new network file can be defined for cases where it is felt that significant changes in the network have taken place.

Rail

In a similar way it is possible to construct a network for rail movements using a simple routing algorithm. Some care is needed to allow for the presence of significant interchange points. The model can be checked against routings suggested by railplanner software and timetable information.

A.11 Congestion modelling

Travel times for the routes specified in the highway network files are calculated using speed/flow equations. Each road type modelled within the STM has it own speed/flow equation.

The model requires various base year speed data for the highway network and parameter values relating to speed/flow equations. The base speeds are therefore an input for the base year. In a forecast year speed/flow equations are applied for each zone to each road type to calculate the speed. The speed/flow equations are functions of the pcu-kms in the base and the forecast cases. These equations are incremental in character in the sense that difference between the base speeds and the forecast speeds is driven by the pcu-kms in the forecast year relative to the base pcu-kms and that the model is guaranteed to give the base speeds when the base model scenario is run.

The speed/flow equations and their calibration are based, in part, on the link speed/flow in COBA but have a more sophisticated form for urban areas. For urban centres the equations are based on earlier studies in which the equations describe the area-wide variation of average speeds as congestion levels change. A minimum speed can be applied at very high traffic levels.

STM also includes a model of crowding on public transport.

A.12 Representation of freight

Within the current STM freight is as a loading on the highway network. Freight levels are treated as given proportions of all traffic.

A. 13 Parking models

The detailed parking model is used to model the impact of parking policies and capacity limitations within a zone. Because the parking model slows down the process of convergence to equilibrium (being relatively more prone to oscillations when compared with other model components) it is not used in all zones. Instead, it is limited to certain "parking zones" which are normally town/city centres and coincide with or at least contain the main parking control areas. Parkers in other zones are treated more simply; each is subjected to a tariff based on the assumed duration of stay for each trip purpose.

Demand for travel by a mode into a parking zone is influenced by the generalised cost (GC) for travel by that mode between the origin zone and the parking zone. GC combines money and time costs into a single measure of money for a given trip by the mode. Both money and time costs are made up of separate component elements. For example, money could be a combination of fares, cordon charges and parking charges. Travel time can be separated into in-vehicle time, waiting time, access and egress time. It is converted to money using an appropriate value of time. Each mode has its own appropriate form of GC equation i.e. is a particular combination of these costs elements. In the case of car, we have

- fuel costs
- time costs
- cordon charges
- parking cost

The parking cost comprises a money cost (the effective tariff) and a component which relates to searching for a parking place. This element becomes more important as capacity is reached. The sum of the money and search cost equals the parking cost (PC) element. As the PC increases, the cost of travel by car into the destination zone increases. This can have two effects: a shift to other modes (typically to public transport) and redistribution to other zones. Redistribution would, for example, encourage movement of work trips to other destinations where jobs were available. This would imply a change in jobs by travellers (a long-term effect). The parking model does not, on the basis of generalised costs alone, cause travellers to maintain their ultimate destination zone, but move the location of their parking outside that zone. There is, however, a mechanism (outlined below) for relocating parking because of insufficient parking capacity.

In addition to calculating the PC the parking model estimates the impact on traffic congestion and travel times due to high parking demand. Normally parkers are assumed to park within the zone in which they ultimately destinate. An exception to this would be P&R when a new P&R is introduced in future years using the new mode facility. This does not apply to the base because no P&R sites are modelled in the base.

Another exception arises when overall parking demand is too high for the total capacity. In this case the parking model determines a proportion of "out-parkers" and assigns to

these a penalty cost (a parameter of the model). Out-parkers are assumed to ultimately destinate in the destination zone (where their activity, e.g. work, is located) but park in an adjacent zone. This is achieved in STM using a simple mechanism in which out-parkers from the destination zone are allocated to pre-defined adjacent zones (out-parking zones); the proportions (parking distribution factors - PDF) allocated to each of these zones are set in a data file. Normally the PDF are not changed over time, but they could be made year-dependent. The PDF are applied to parkers independent of purpose, origin and household car ownership. The same factors are used in both the am peak and the interpeak periods.

In the DISTILLATE STM we have used the new parking model developed originally for the West Yorkshire STM for their LTP work. This model is able to represent parking demand and supply for

- PNR
- Long stay pay
- Long stay free
- Short stay pay
- Short stay free

The user can set policies for changes in the supply of these parking categories and the charges for them. PNR parking can also have charges applied.

The model starts from a base picture in which demand is defined across the parking supply types for each category of trip purpose and where the supply for each parking type is estimated. In a forecast year parking conditions may change as a result of changes in demand (e.g. due to increased retail opportunities, jobs and car ownership growth) and policies, such as increased tariffs and reductions in parking provision. These changes are accommodated in the model using a series of simple algorithms which, for example, distribute rising demand over the available spaces in sequence dependent on the types of places available. When supply is exhausted cars are forced to redistribute as described above.

A. 14 Equilibrium modelling and convergence testing

The main loop of the STM model (see Figure B1) balances transport supply and demand. Generalised costs for travel between zones determine the mode and destination choice of trips for each car ownership category, purpose and time of day. The new pattern of trips then leads to a new set of generalised costs through the application of the congestion model and capacity constraints. STM generalised costs are responsive to road capacity constraints (which affect speeds and thus travel times), public transport crowding levels (affecting perceived value of time for transit) and parking demand (affecting parking search time, access and egress costs, and additional traffic capacity effects). Many other components of costs are also modelled, but it is the supply-side costs that primarily influence the balance in the damped feedback loop. Convergence to equilibrium has been reached when measures of capacity (such as speeds, crowding, mean parking charge) are consistent within a specified tolerance level between successive iterations. The principal test is for the convergence of speeds. A number of damping calculations are also made to reduce oscillations in the convergence process.

A. 15 Calibration and validation of the transport model

Calibration of STM is terms of published elasticities for the responses of travellers to various costs factors. Parameters controlling modal split are adjusted to ensure that for trips within the study area the levels of response in the model match those observed for public transports and private car travel. In effect, the calibration determines the response of demand to generalised cost by changing certain elements in the generalised cost expression (e.g. bus fares). The 'generalised costs elasticity' in conjunction with the weight factors for the cost elements (most importantly value of time) then determines the response for those elements.

Validation of STM means comparison of STM outputs against expectation. 'Expectation' can mean

- known or expected forms of behaviour within certain ranges of response
- independent data (i.e. not used in setting up the model) such cordon counts, modal shares etc.

In addition, a model can be validated by performing a series of realism checks when specific policies are applied.

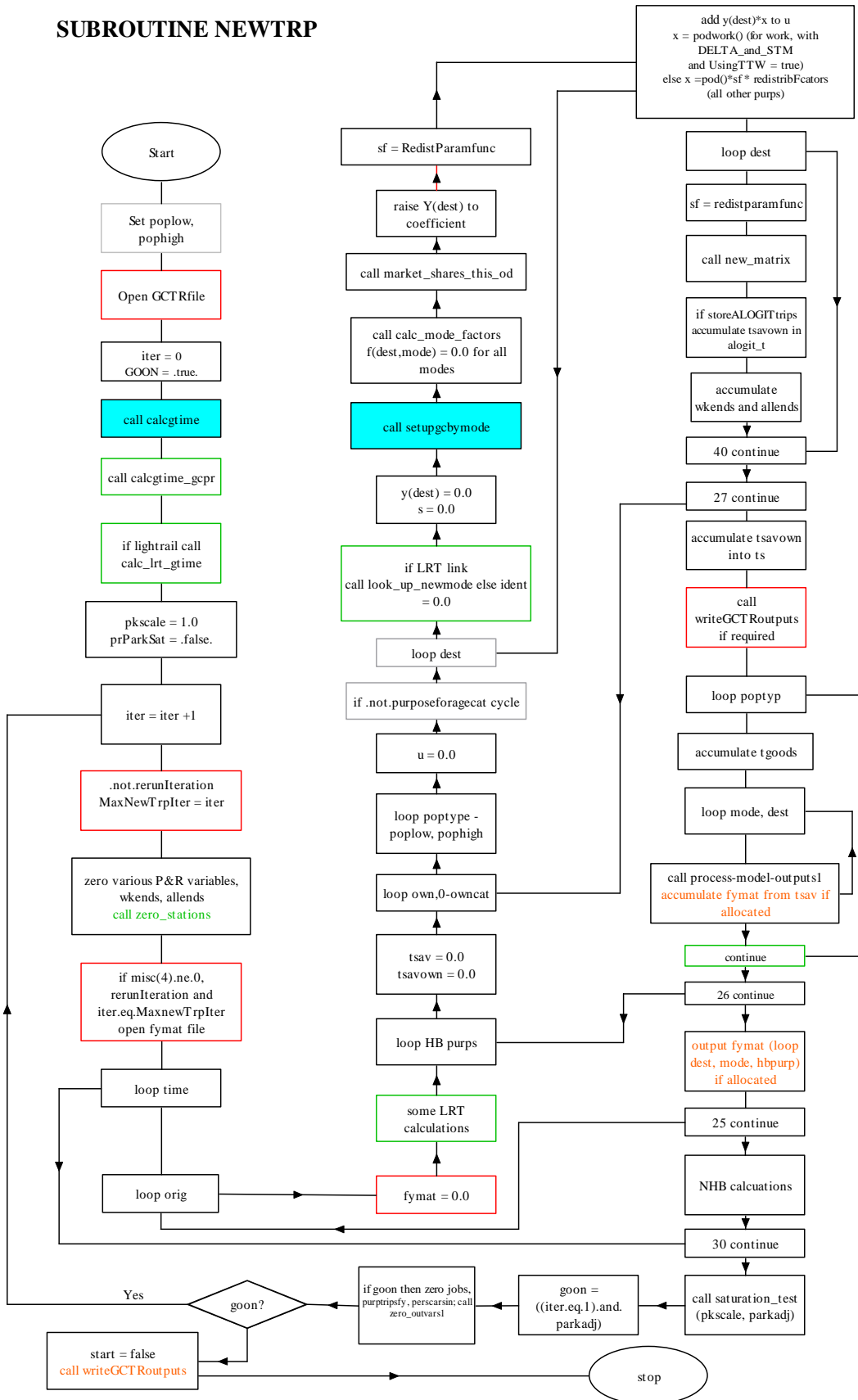# Appendix B: Flowchart of the main demand module in

SUBROUTINE NEWTRP

Figure B1.