Transportmetrica, Vol. 3, No. 1 (2007), 1-19

FORECASTING TRAVEL FOR VERY LARGE CITIES: CHALLENGES AND OPPORTUNITIES FOR CHINA

DAVID BOYCE[1] AND CHERRY XIONG[2]

Received 27 August 2006; received in revised form 27 October 2006; accepted 27 October 2006

Urban travel forecasting, a component of the field of urban transportation planning, has entered its sixth decade, having been introduced in the early 1950s in North America, from which it spread to other developed countries, and then to developing countries such as China. Although this field is relatively mature, many new research directions are being pursued; at the same time practitioners experience ongoing difficulties with applying commercial software systems in correct and meaningful ways. China is presently undergoing rapid urbanization, implying major investments in road and transit systems, as well as airports and intercity surface transportation systems. Wise investment decisions require forecasts of future urban travel in China's congested cities. To understand the state of travel forecasting, the authors recently held discussions with professionals at several planning agencies of China's largest cities, as well as university departments engaged in related teaching and research. Our objective here is to relate the general findings and impressions from these discussions, as well as the message we presented to the seminar participants.

KEYWORDS: Urban travel forecasting, congestion, transportation planning

1. INTRODUCTION

Urban transportation planning agencies in China have engaged in travel forecasting, in the sense that term is used in North American and European practice, for nearly 20 years. (Planning agencies in Hong Kong and Taiwan have undertaken travel forecasts for a longer period, but are not considered in this paper.) As one benchmark for this statement, we understand that western travel forecasting software was introduced into China as early as 1988. Moreover, the principal international travel forecasting software firms all have a definite presence in the China market today.

In order to understand better the state of travel forecasting, the authors recently met with professionals at several planning agencies of China's largest cities, as well as university departments engaged in related teaching and research. Our objective here is to relate the general findings and impressions from these discussions, as well as our message to the participants.

A summary of the seminar offered to transportation planning agencies during our meetings in June 2006 constitutes the first part of this paper. Then, our observations on the state of practice in China, as gleaned from seminar participants, are summarized. Further discussions with agency professionals led us to think anew about the role of travel forecasting in rapidly developing cities, which almost by definition have rather poor data about the current situation and the near term future. In this section we offer some preliminary notions concerning research problems that may be pertinent for China's cities and other rapidly developing cities. We conclude the paper with our thoughts about the future prospects for our field in this exciting environment.

[1] Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL 60208 USA. Corresponding author (E-mail: d-boyce@northwestern.edu).
[2] PTV America, Inc., 1300 N. Market Street, Suite 603, Wilmington, DE 19801 USA.

2. HISTORY AND STATUS OF THE FIELD OF URBAN TRAVEL FORECASTING

Transportation planning for large cities as it is known today began in America in Detroit in 1953, building on previous experience with travel surveys in the 1920s and following World War II. The Detroit study approach was transferred to Chicago in 1955, where it evolved substantially. Concurrently, the Bureau of Public Roads of the U. S. Department of Commerce initiated studies and development of methods in support of these and other American urban transportation studies. From these initial experiences, our field spread over North America, to Europe, Australia-New Zealand and Asia.

Early travel forecasting for the evaluation of alternative road and transit plans was based on large-scale household surveys, relatively simple methods and very early main-frame computers and associated machinery suited to handling large data bases. Road network design methods were also devised and applied in the Chicago study, but without lasting success. By the 1960s methods for forecasting land development were devised, with strong relationships to travel conditions and costs. Essentially all of these methodological developments occurred within these studies with few contributions from universities or other research organizations. Likewise, university courses related to this field were based on the study reports, the most prominent being the Final Report of the Chicago Area Transportation Study and Creighton (1970) based upon the Chicago study and his subsequent experience.

Slowly, government laboratories and universities began to contribute models and methods to the field. One of the first was by Wardrop (1952), a scientist at Britain’s Road Research Laboratory, which included a brief description of the concept of user-equilibrium route choice: all used routes between an origin and a destination have equal and minimal travel times. At almost the same time, a theoretical model relating origin-destination travel and route choice to user-equilibrium travel costs over a road network was proposed in 1954 by Beckmann et al. (1956), then at the University of Chicago. A model describing the spatial relationship of households, jobs and shopping/service activities based on travel cost relationships was proposed by Lowry (1964) at the University of Pittsburgh. From the late 1960s onwards, methods for solving large-scale travel and location models were devised by several researchers, many of whom were Ph.D. students in American, Canadian and British universities. Even so, many years often passed before these contributions were implemented into professional practice.

The use of main-frame computers was central to early transportation and land use planning; indeed, the early transportation studies were among the first civilian uses of main-frame computers for problem-solving. Government agencies in America and Britain invested heavily in software development from 1955 to 1980. Computer manufacturers began to develop software systems in the 1960s as a way to sell computer time to planning agencies. Later, the activities of commercial software developers intensified with the advent of personal computers and engineering workstations in the early 1980s. During the past 20 years, software systems have matured, and today several international firms dominate this market (Caliper, Citilabs, INRO, PTV).

What, then, is the status of our field today at the international level? Software systems continue to improve in speed and rigor of the methods represented. Traffic simulation systems of various types and approaches are presently available or under testing, including dynamic traffic assignment models. University researchers and consultants are urging the use of activity-based and tour-based travel models, in place of the traditional trip-based models. Basic travel forecasting practice, however, lags far behind the state of the art in two ways: 1) traditional methods devised long before model properties were well understood are still in use; 2) solution techniques are improperly applied without an appreciation of their properties.

Even in countries with much experience with these methods, it is sometimes not recognized that software systems are basically toolkits for implementing a model by the professional travel forecaster. Hence, these software systems are generally not actually models themselves. Moreover, software developers are constrained to a substantial extent by what practitioners believe they want and need in a software system, even though their desired improvements may actually be incorrect. Finally, road and transit network design methods, as were proposed in Chicago in the late 1950s, have completely disappeared from use.

What are the implications of this situation for China? Is it a dilemma, in the sense that useful methods are not sufficiently applied, or does it present opportunities? Our sense is that travel forecasters in China have opportunities to avoid the errors and shortcomings of traditional practice and instruction by jumping over mistakes of past generations of practitioners. They also have opportunities to contribute new approaches, unconstrained by conventional thinking of past practice. Their dilemma, if any, is one of developing new solutions to difficult problems, where others have failed. And the challenges they face include acquiring mathematical and scientific expertise in fields that has long been regarded as lacking rigorous and sophisticated approaches.

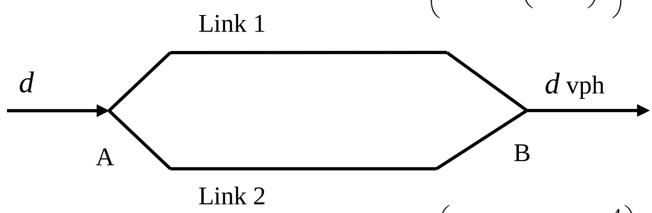
In the next two sections we explore some dimensions of these problems through two case studies based on present-day American practice. Many topics could be chosen for discussion here. We have selected two that we believe are fundamental to the quality of travel forecasts in the presence of heavy traffic congestion. We believe these issues are important, and perhaps critical, for the evaluation of alternative plans or scenarios in China. The first concerns Traffic Assignment; the second examines the solution of the Sequential Travel Forecasting Procedure with Feedback. Both case studies focus on questions of the convergence of the model solution, a relatively technical property of these models, and are closely related to the performance of commercial software systems.

3. CONVERGENCE OF TRAFFIC ASSIGNMENTS

Traffic Assignment is the traditional term given to the problem of forecasting route flows and link flows in a road, or transit, network, given the zone-to-zone flow of travelers for a specified period such as the morning peak period. Solving this problem well is especially critical for cities with highly congested road networks. The model we consider here is known nowadays as static traffic assignment, since it considers the flows to be constant over the period of analysis, in contrast to dynamic traffic assignments in which flows for short intervals of perhaps five minutes interact over the network from interval to interval.

The Traffic Assignment Problem was first formulated and analyzed by Beckmann et al. (1956), but their findings were unknown to practitioners for at least ten years. Practical algorithms for solving the problem for large networks were proposed in the early 1970s, implemented in practitioner software systems by the end of the 1970s, and adopted as the method of choice over earlier heuristic solution methods. The principle of the problem formulation may be readily grasped by considering a two-link network, shown as Figure 1. By plotting the travel time functions from opposite ends of a chart with an assumed flow of 4,000 vehicles per hour (vph), the solution at which Wardrop's principle of equal travel times on each link is satisfied is readily identified as the point where the functions intersect (1,522, 2,478), as shown in Figure 2. Less apparent is the fact that this intersection corresponds to the solution that minimizes the area under the two curves. This result can be observed by considering another solution with (1,522 – *k*) and (2,478 + *k*). In this case, the area is increased by the amount of the triangle lying between the two functions and the vertical line at (1,522 – *k*). As is widely understood, this problem can be formulated as a convex optimization problem with the conservation of flow constraint, $f_1 + f_2 = d$, and non-negativity constraints, $f_1, f_2 \geq 0$.

$$c_1(f_1) = 15 \cdot \left(1 + 0.15 \cdot \left(\frac{f_1}{1{,}000}\right)^4\right)$$



$$c_2(f_2) = 20 \cdot \left(1 + 0.15 \cdot \left(\frac{f_2}{2{,}000}\right)^4\right)$$

FIGURE 1: A two-link network with fixed flow

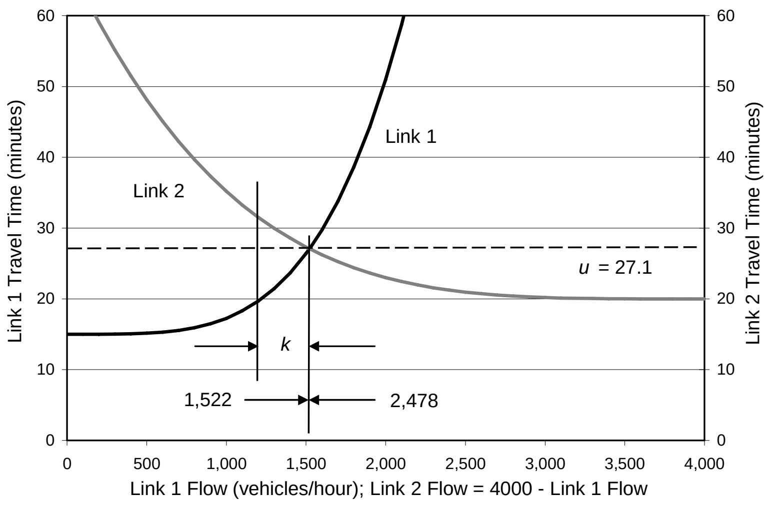


FIGURE 2: Solution of a two-link user-equilibrium traffic assignment problem

Although large problems could be solved approximately by the end of the 1970s, they could not be solved precisely with reasonable amounts of computational effort because the known algorithms converged only very slowly after the first several iterations. Such solutions were considered adequate then, and that attitude tends to prevail even today. From time to time, however, one hears accounts of practitioners who are dissatisfied with their solutions because link flows far removed from a proposal under study are different in the scenarios under consideration, such as No-Build versus Build scenarios.

This situation may cause travel forecasters to lose confidence in their results, and in any event illustrates that the solution has not satisfactorily converged.

Such a case came to our attention in 2001, which led to a small study by Boyce et al. (2004). Transportation planners at the Delaware Valley Regional Planning Commission were asked to evaluate the effect of adding a pair of ramps between two freeways, east-west I-295 and north-south NJ 42. These freeways are shown in Figure 3 as the two heavily used facilities crossing in the lower center of the figure. When traffic assignments with and without the ramps were compared, they found that 24-hour link flows had changed all over the 40,000 link network covering a portions of southeastern Pennsylvania and southwestern New Jersey. Upon examining maps of their results, the apparent changes in link flows were clearly caused by the inadequate convergence of their traffic assignments.

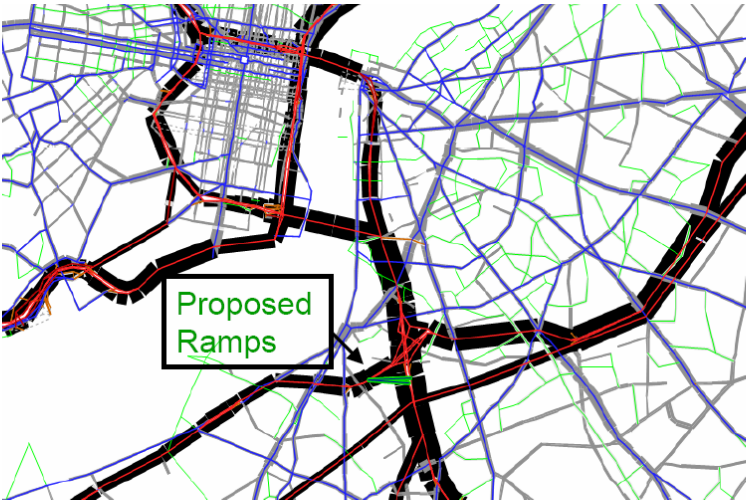


FIGURE 3: Location of the proposed ramps connecting I-295 and NJ 42

Traffic assignments were prepared for the road network with and without the ramps, and the differences in link flows were computed for a range of levels of convergence. The convergence measure used in this case was the Relative Gap expressed as a percentage. At convergence, the Relative Gap is zero, signifying that all used routes have equal and minimal travel times. The Gap, also known as Total Excess Time, is the difference between the travel time over each used route and the corresponding minimum time route, weighted by the used route's flow, and summed over all used routes in the network. The Relative Gap equals the Gap divided by the sum of the areas under the link cost functions, which is minimized to obtain the equilibrium solution, as illustrated in Figure 2. The algorithm used for these two solutions was the Origin-based Assignment (OBA) algorithm of Bar-Gera (2002), which is known to converge to the precision of personal computers with 32 bit word length. In this case the Relative Gap is no larger than 10^{-11}%. Link flow differences (flow with the ramps less flow without the ramps) in the vicinity of the ramps were plotted versus the Relative Gap, as shown in Figure 4. As the figure shows, the flow differences for these links changed substantially as the two assignments converged. A Relative Gap of 1% generally corresponds to best current practice, and is roughly equivalent to 25 iterations of a widely used assignment algorithm known as the Frank-Wolfe method. Numbers of iterations for more highly converged assignments are also shown on the figure; see Boyce et al. (2004) for more details.

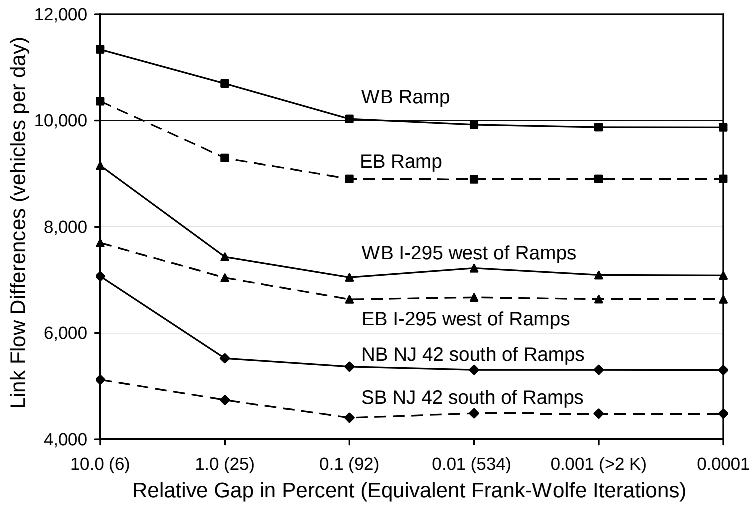


FIGURE 4: Link flow differences for selected links in the vicinity of the proposed ramps

Maps of link flow differences throughout the network were also prepared, and compared with the link flow differences from very highly converged solutions. The difference in the link flow difference at a given Relative Gap and a highly converged solution is termed the Error in Link Flow Difference (Error in LFD). These errors are shown in Figures 5 and 6 for two levels of convergence, Relative Gaps of 1% and 0.0001%, for the 8,000 links nearest the pair of ramps. These maps show that at a Relative Gap of 1%, many links have errors ranging up to +/-1,000 vehicles per day. With a Relative Gap of 0.0001%, the errors are reduced to +/-10 vehicles per day. Similar results were observed for maps of the entire region.

As can be seen from these two figures, which are representative of more detailed studies, achieving adequate convergence levels for assignments is the key to obtaining sensible and useful estimates of traffic flows. Fortunately, some developers of travel forecasting software are beginning to incorporate results of recent advances in traffic assignment algorithms into their software systems. As these methods become available, practitioners ought to perform their own computational experiments to determine what levels of convergence are most meaningful for their travel forecasting requirements.

4. SOLVING THE SEQUENTIAL PROCEDURE WITH FEEDBACK

As suggested in the above historical review, the problem of predicting origin-destination travel and the associated route choices on a network was formulated and analyzed as a single problem at the outset of our field. Practitioners were unaware of this formulation, however, and devised their own method of forecasting travel, known as the four-step or sequential procedure. Actually, this procedure can be seen to be a rough heuristic for solving the formulation of Beckmann et al. (1956) and later specializations of it. Once established in early transportation studies during the 1955-1970 period, this procedure took on a status that was rarely questioned. Ignored during subsequent years was the lack of consistency between the travel time assumptions forming the inputs to this procedure and the resulting travel times determined by solving the traffic assignment step. In 1991 a regulation was issued by the Federal Highway Administration in America requiring that the sequential procedure be solved "with feedback". The dilemma presented by this requirement was that practitioners did not know how to solve the procedure with feedback. Moreover, research findings on the formulation and solution of an integrated version of the problem were largely unknown.

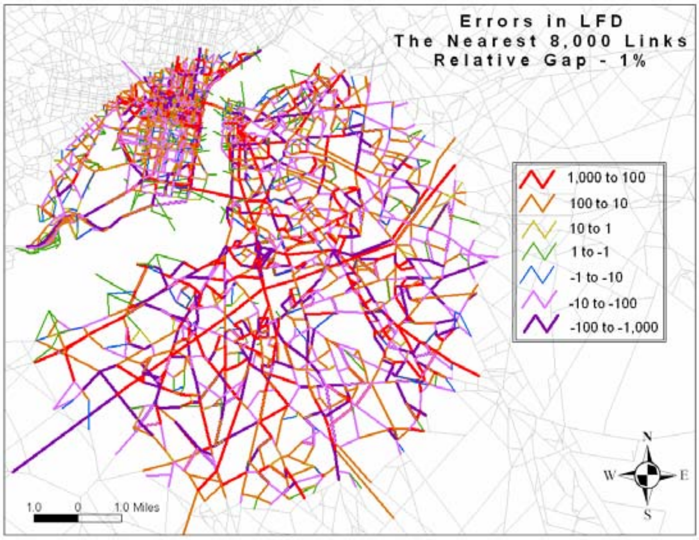


FIGURE 5: Errors in link flow differences at 1% Relative Gap

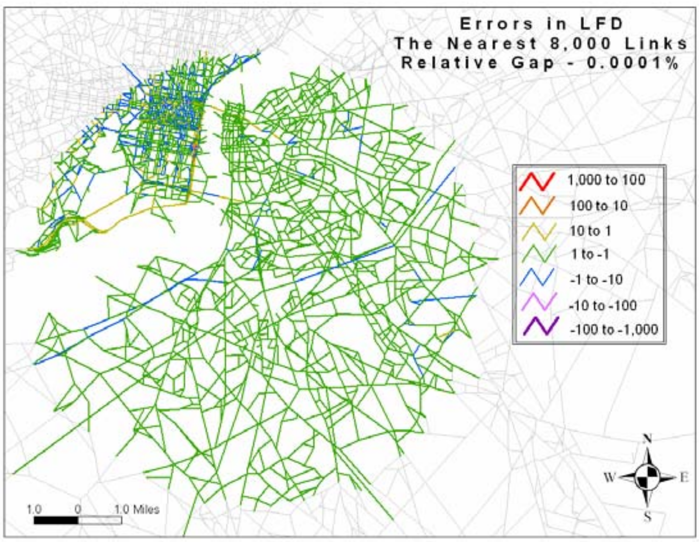


FIGURE 6: Errors in link flow differences at 0.0001% Relative Gap

University research did continue during the 1980-2000 period, however, building on the findings of graduate student research in the early 1970s. Most prominent among these was the formulation and solution algorithm proposed by Evans (1976). Her findings provided the basis for research by Lam and Huang (1992a,b) for Hong Kong, and Boyce (1984), Boyce and Zhang (1997) and Boyce and Bar-Gera (2003) for Chicago. Moreover, Bar-Gera and Boyce (2003) showed that these problems could be solved more efficiently with Bar-Gera's OBA algorithm; see also Boyce (2002). Since these papers are generally available, we will not review them further here.

The above research findings for integrated models also provided insights on how to solve the traditional sequential procedure with feedback. Contrary to intuitive notions of simply using the link travel times determined by the most recent assignment to construct a new table of zone-to-zone travel times, and re-solving the four-step procedure, Boyce and Bar-Gera (2006) argued that averaging of successive trip tables is necessary. In the following, we refer to procedures that do not involve averaging of the trip tables as a naïve feedback method. One version of the averaging procedure is illustrated in the flowchart shown as Figure 7.

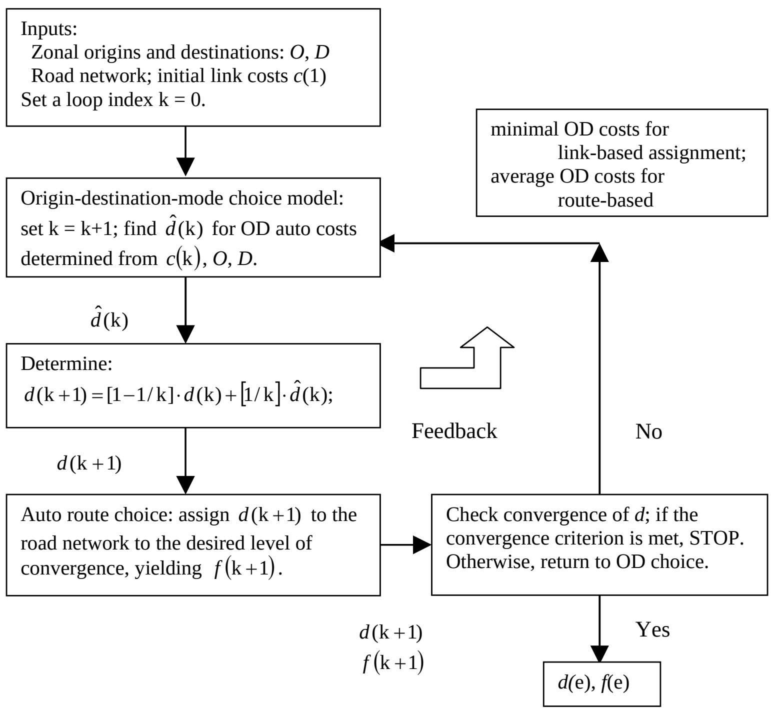


FIGURE 7: Flowchart for solving the sequential procedure with feedback

Two methods of averaging of the current trip table d with a newly computed table $\hat{d}$ are evident. One is the Method of Successive Averages (MSA), in which the weight associated with each new trip table $\hat{d}$ is equal to the reciprocal of the iteration number, as shown in Figure 7. MSA, introduced into transportation modeling by Powell and Sheffi (1982), converges to the equilibrium solution under very general conditions, although convergence may be quite slow. The other possibility is to use a constant weight, such as equal weights on the new trip table and the previously averaged table. Following the averaging procedure, the resulting table is assigned to the road and transit networks. If a link-based procedure is used, then an entirely new assignment is required, since link-based assignments cannot be updated. If a route-based or origin-based assignment algorithm is applied, then the previous assignment can be updated to reflect the changes to the trip table; these algorithms can also determine the average OD costs, which are more suitable than the minimum OD costs for computing the next trip table. For link-based methods, however, only the minimum cost routes are available.

Results for the Naïve, Constant Weight and MSA Feedback methods are shown below for a 500 zone model of the Capital District, Albany, NY, implemented in the PTV software system VISUM. The model has five classes of travelers with a total evening peak period OD flow of 313,082 persons per hour. This comparison uses Total Misplaced OD Flow (TMF) as a convergence measure. Total Misplaced OD Flow is simply the sum of the absolute values of the differences in cell values of d and $\hat{d}$ over all cells in the trip table. When the two tables are equal, then TMF equals zero. Intuitively, the feedback procedure has then converged: a trip table has been found that when assigned to the network yields zone-to-zone travel times that can be used to compute a new trip table with exactly these same values. Satisfying this condition is the essence of solving the Sequential Procedure with Feedback.

Total Misplaced OD Flows for the three solutions are shown in Figure 8 using the log scale to facilitate their comparison. For Naïve Feedback, TMF reached a minimum value of 34,200 persons/hour in the 7th loop, which shows a large discrepancy between the trip table in the previous feedback loop and the trip table computed based on the travel times resulting from assigning that table in the current loop. At full convergence, TMF must equal zero. A value of 0 persons/hour was achieved by the Constant Weight (CW) method in the 17th loop, and a value of less than 100 was reached in the 8th loop, using weights of 0.6 for d and 0.4 for $\hat{d}$.

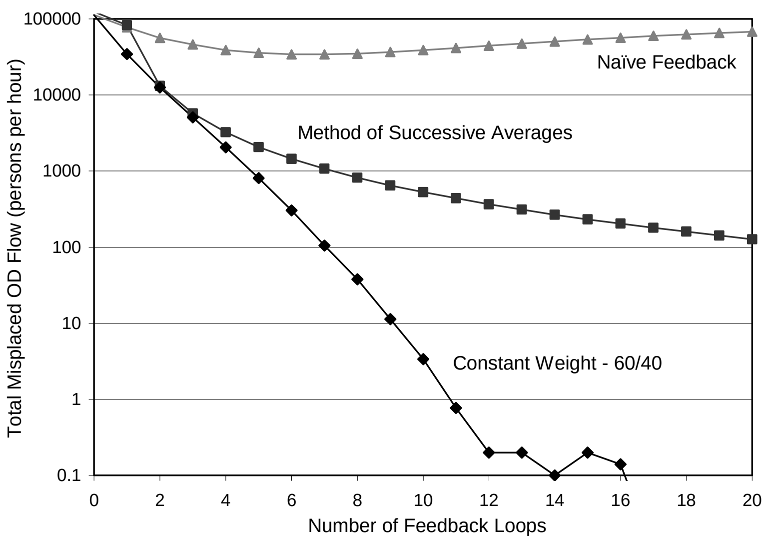


FIGURE 8: Convergence of the feedback procedure based on Total Misplaced OD Flow

The MSA method reached a minimum TMF of 127 persons/hour in the 20th loop when the process was terminated. The values are acceptably small for the MSA and CW methods.

The number of assignment iterations in each loop required to achieve a Relative Gap of less than 0.01% is shown in Figure 9. For comparison with the first case study, the convergence measure shown is the Relative Gap, computed for the traffic assignment only. For Naïve Feedback, the number of assignment iterations falls to 12 in loop 6, but then increases gradually to 27. For the other two methods, the number of assignment iterations falls rapidly to the minimum number of 2. In the route-based assignment algorithm applied in VISUM, the assignment is updated following the revision of the OD table. If a link-based method were used, then a new assignment solution would be required in each feedback loop.

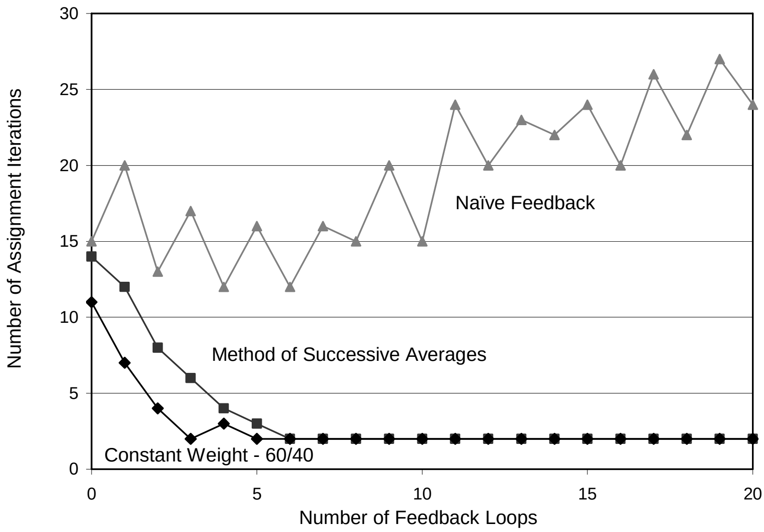


FIGURE 9: Number of assignment iterations required for a Relative Gap less than 0.01%

These results, as well as the extensive studies of the convergence of integrated models cited above, show that forecasts of origin-destination flows and link flows depend upon solving the sequential procedure properly with respect to travel times. In congested cities with several interacting modes, representing the effects of congestion properly is bound to affect the forecast of origin-destination flows and link flows in a profound manner. In less congested cities, as found in some developed countries and in smaller and more dispersed cities, the effects of congestion may not be so important.

5. OBSERVATIONS ON TRAVEL FORECASTING FOR CHINA'S LARGE CITIES

The first author has visited urban transportation planning agencies throughout his career in an ongoing effort to understand the requirements for travel forecasting models, and to assess the state of practice; see Boyce et al. (1970) as one example of such an effort. From this experience he is well aware that spending a few hours in a planning agency does not enable one to draw definite conclusions about the modeling activities underway there. Nevertheless, it seems important to summarize what we learned in visits with travel forecasters in two of China's largest cities, if only as a basis for further discussions and dialogue, as well as to encourage comments on our observations.

We were advised repeatedly that travel forecasters working in China's large cities face difficulties with regard to two important factors. First, the cities are expanding very rapidly, including reconstruction of large areas that were historically developed at relatively low densities (one and two story dwellings in compact street layouts). Second, data are relatively scarce, and available data are often not shared among planning agencies. The second problem is also exacerbated by the first, since travel surveys and network data in rapidly changing areas may be regarded as a waste of time and resources. Both of these problems raise interesting questions about the way in which planners organize and perform their work.

In thinking about travel forecasting for large cities undergoing major investments in road and transit systems, and experiencing rapid economic development and associated motorization, we have focused on two questions. First, do the travel forecasting methods adequately represent the congested traffic conditions found in these cities? In other words, does the presence of substantial traffic congestion lie at the heart of the travel forecasting procedure? Second, are shortcomings of travel forecasting practice with regard to highly congested systems recognized by China's modelers, and are steps being taken to address them? That is, are China's travel forecasters learning from the mistakes of their predecessors in other countries, or are they doomed to repeat them?

While there are many ways to organize this discussion, we have elected to proceed according to the traditional sequential procedure, including the question of the consistency of the solution, or solving the procedure with feedback. Then we turn briefly to a few other questions that were raised in our discussions.

5.1 Trip generation

In view of the limitations of data for base year conditions, and the rapidly changing urban landscape, estimation of the amount of travel leaving and entering small areas (zones) of cities is itself problematic. We did not explore this problem during our discussions, but we understand that urban development of rural areas and redevelopment of previously urbanized land in China is subject to complexities that extend well beyond the scope of this paper. What seems clear is that the experience of planning agencies in both North America and Europe with land development models is not applicable to the present situation in China.

Probably, the best that can be done in China in the near term is to make assumptions of future zonal activity levels based on construction plans and development trends, or on assumptions of the planning agency about future development, and to use trip generation rates derived from recent travel surveys to estimate zonal travel departures and arrivals. The uncertainties associated with these levels are clearly challenging, and are addressed further in Section 6.

5.2 Trip distribution

In our discussions, approaches to preparing trip tables for future years raised concerns with respect to their validity as a forecast. The approaches described are also found in some American and European practices, about which we also have concerns. One procedure described is first to estimate a base year trip table with survey data and a trip distribution model, taking into account interzonal travel times. Then, the base year table is adjusted (factored) to the forecast year on the basis of changes in originations and terminations resulting from land development and redevelopment by zone. We understand this trip table pertains to travel during a typical 24 hour weekday. This trip table is then held fixed in conducting mode choice and traffic assignments for alternative road and transit plans.

Precedents for this factoring procedure are found in early American travel forecasting methods, and may still exist in some practices. Clearly, such a trip table has no definite relationship to future congested road and transit network travel times or monetary costs. Therefore, the validity of this approach is questionable, if the result is intended to describe the behavior of travelers.

At the same time, we note that the use of a fixed trip table is required in certain American transit planning procedures. In this case trip tables are required to be held fixed to prevent the transit forecast from being biased in favor of a preferred modal or investment alternative. For example, cases have been observed in which forecasts for a rail transit alternative were higher than those for bus services with similar service attributes (Pickerell, 1989). We also understand that fixed trip tables for a future year are sometimes used in European practice, but we are uninformed about the rationale.

We also strongly question the use of 24-hour trip table forecasts, especially if such forecasts are converted to peak period forecasts by factoring using base year travel surveys or link counts. Regardless of its basis, a 24 hour trip table cannot be functionally related to congested travel times, as those travel times change very substantially over the 24-hour weekday. Moreover, the use of 24-hour roadway capacities equal to factored hourly capacities is highly questionable. The use of 24-hour trip distribution models originated in the earliest American urban transportation studies. Such a shortcut was reasonable then in view of the severe limitations of computer size and work schedule faced by early travel forecasters. Moreover, in those early models, a representation of the changing level of service over the 24-hour day in the model was included. Present day models typically include no such representation.

Finally, future trip tables should be based on the travel times and costs of all modes. A sound basis for representing the cost of all modes in a trip distribution model is the composite cost function (Williams, 1977), also known as the inclusive value. Desirably, such a function should reflect the travel times of bicycle and walking modes, as well as car and public transit, in an appropriately structured nested mode choice function.

5.3 Mode choice

That mode choice is described as the most difficult problem facing travel forecasters in China’s large cities is not unexpected. As a point of reference, the modal distribution (excluding walking) for all trip purposes in Beijing was estimated for 2004 based on travel surveys as follows: bus and metro – 30%; bicycle, motorized bicycle and mopeds – 40%; private cars, and cars provided by government agencies and firms – 20%; and taxis – 10%. Forecasting these mode choices for the future is clearly a very substantial challenge.

Modal transportation services in China’s large cities are changing very rapidly. Several metro lines are presently under construction in both Beijing and Shanghai, and bus rapid transit is being introduced in Beijing; these improvements will encourage travelers to shift to transit from other modes, but perhaps especially from bicycle and moped. As motorization is one of China’s economic development strategies, families are encouraged to purchase a car. In Shanghai, however, a car ownership restraint policy has been introduced by raising the cost of car ownership through licensing fees. The timing of these developments is likely to confound the efforts of travel forecasters for some time to come, especially given the uncertain timing of transit improvements and increasing rates of car ownership.

5.4 Traffic assignment

Representation of congested travel times involving a mixture of modes operating on roads (buses, cars, various types of bicycles and mopeds, and pedestrians) is a major challenge for traffic assignment models. In the absence of travel time-flow (volume-delay) functions for mixed modes, traffic assignment practice may be limited to estimating the future use of roads that are restricted to cars, trucks and buses. The solution of the traffic assignment problem with a mixture of modes in past practice was based on the use of equivalency factors to convert trucks and buses into car equivalents. This approach may be unsuitable for China's large cities because of the diversity of vehicle types and their operating characteristics. For an example of an operational model incorporating several vehicle types, see Wu et al. (2006). More generally, travel time-flow functions for each of the several modes operating on arterial streets (cars, taxis, buses, cycles and pedestrians) are needed. The resulting problem is technically challenging, but represents an opportunity for China's well-qualified transportation academics, as well as its professionals, to contribute to the state of the art.

The standard traffic assignment model used throughout the world is also deficient with regard to the way that travel times are said to depend only on a link's own flow, and not on conflicting and opposing flows at downstream intersections. The former functions are termed separable. The use of non-separable travel delay functions poses a difficult technical problem, one that has been ignored for too long by travel forecasting professionals in developed countries. The solution of the traffic assignment problem with non-separable delay functions is important for the large, congested cities of China. Again, this is a problem area in which Chinese academics could be working actively.

5.5 Feedback

Solution of the four-step procedure with feedback is undertaken in some planning agencies of China's largest cities, but as in North American travel forecasting practice, solutions with feedback may not be implemented effectively. In particular, the ineffective procedure described in Case Study 2 as Naïve Feedback is used in some situations. As a result, and in view of the practice of using a fixed future year trip table, the extent to which congested conditions are consistently represented in forecasts is problematic. We believe that solving the four-step procedure properly with feedback, or alternatively solving an integrated model that assures consistency between (1) the travel times and costs on which the trip table is based, and (2) those that are determined by the assignment of that trip table to the network, is a matter of high priority for China's travel forecasters, if they intend to represent traffic congestion in their forecasts.

Travel forecasters in Chile have tackled this problem differently than others. The software system ESTRAUS was designed to solve an integrated model and was applied in restructuring the bus network of Santiago, Chile; see http://www.mctsoft.com/, Fernandez et al. (2005) and also Siegel et al. (2006). This example offers evidence that substantial progress in solving integrated models of travel and route choices for the study of congested urban transportation networks has been achieved and is proving to be useful in practice. In contrast to software systems designed as toolboxes for use in solving the sequential procedure and related tasks, ESTRAUS was designed to solve a specific class of models, about which practitioners may not be so well informed. Moreover, solving the same integrated models with the toolbox systems may pose serious difficulties to the typical practitioner who may not be well-versed in the technicalities of integrated model solution algorithms. Although this approach has not been embraced by other practitioners, it could be useful and beneficial to emerging travel forecasting practice in China and other rapidly developing countries.

5.6 Travel forecasting software systems

China's transportation planning agencies and institutes generally have obtained present-day travel forecasting software through system developers in North America and Europe. Generally, however, these procedures do not appear to be applied with an adequate understanding of model attributes and solution properties. While such software can be very helpful to the well-trained and experienced travel forecaster, in most cases the software itself does not constitute a travel forecasting model. Software companies do offer training courses on the use of their software systems; some companies also provide detailed manuals on how to implement, calibrate and solve models. In some cases these manuals are also the basis for short courses on modeling and forecasting.

5.7 Road and transit system design

In response to queries about road system design, ring road concepts and experimentation with alternative system layouts across China's many large cities, we noted some interest, but little indication of new approaches. For an example of an early attempt at road system design, see Boyce (2006). No systematic analyses of alternative road system designs were described. In one case we were told that designs of new ring roads tend to follow municipal boundaries and simply serve to expand the existing arterial system, rather than considering whether to restructure the road system in some functional manner.

6. METHODS FOR SCENARIO ANALYSES FOR TRANSPORTATION PLANNING

A point made repeatedly in our discussions with travel forecasters concerns the lack of data on urban activities, and the very rapid change in the level of urban activities over calendar time. In this situation there is a definite tendency to wonder if travel forecasting is worthwhile. This situation can certainly lead to a degradation of the regard for forecasts. By redefining the reason to perform forecasts, however, an interesting and challenging research problem, and analytical approach, can be defined. This section seeks to offer an initial statement of this problem.

6.1 Problem statement

Generally, travel forecasts for a future design year or intermediate year are based on assumptions of the available transportation network capacity and operating characteristics. The performance of the assumed network is then assessed using various performance measures, as discussed below. These measures depend upon point assumptions about zonal activity levels, represented in terms of trip origins and destinations, assumptions that are bound to contain some errors and uncertainties.

Instead of a point assumption, suppose a range of values is assumed for each model input value. The lower end of the range could be the existing activity level, or an estimate of it. The upper value could be an assumption about the likely final development at higher densities, for example. Similar assumptions might be made about link capacities, if appropriate, and even parameter values. Given this set of assumptions, the travel forecasting model can be solved by sampling from the range of values assumed for each input datum, or parameter. Generally, the range defines a uniform distribution, but other assumptions are also possible. If this sampling procedure is repeated perhaps 100 times, resulting in 100 model solutions, then the frequency distribution of summary performance measures of the model output can be viewed as functions of the assumed input distributions.

The availability of these distributions may enhance the types of questions that can be addressed. For example, are the distributions of the performance measures stable, and may they be regarded as “good outcomes” for the proposed transportation network? If so, the proposed network may be regarded as “robust” with respect to the unknown future values of the model inputs. If not, and some performance measures are definitely poor or undesirable, then an attempt might be made to find a more robust solution. Thus, this approach may be used to search for network configurations that are most suitable for the ranges of input data assumed.

A related set of assumptions concerns the staging of improvements. There is often uncertainty about when development, or redevelopment, may occur. For this purpose, sequences of forecasts ought to be prepared for several points in time, such as ten-year intervals including the current or base year. The relation among these cross-sectional analyses, and the suitability or robustness, of plans ought to be examined.

6.2 Performance measures

Measures of transportation network performance for a given spatial distribution of urban activities are reasonably well established. This subsection briefly reviews these measures.

1. regionwide measures
   - vehicle-kilometers of travel by major roadway and vehicle type (e.g. expressway, arterial)
   - vehicle-hours of travel by major roadway and vehicle type
   - space-mean-speed by roadway type
   - mean volume/capacity ratio by roadway type
   - person-hours of travel by mode and user class
   - modal split – proportion of person-trips by mode
   - trip length (distance) frequency distributions by user class
   - travel time frequency distributions by user class
2. subregional measures
   - travel time frequency distributions by origin and destination districts and user class
   - trip length (distance) frequency distributions by origin and destination districts and class
   - mode split by origin and destination districts and user class

6.3 Research methods and issues

Two types of research methods in the literature may be applicable to such scenario analyses. One is stochastic programming, a framework for modeling optimization or equilibrium problems that involve uncertainty. Whereas deterministic optimization problems are formulated with known parameters, real world problems almost invariably include some unknown parameters. When the parameters are known only within certain bounds, one approach to tackling such problems is called robust optimization. Another is chance-constrained programming.

Another approach is the statistical experiment approach of Sacks et al. (1989), which was devised to analyze repeated measurement experimental data. This method was successfully applied by Boyce and Bar-Gera (2003). Boyce and Cote (1966) outlined a similar approach. To pursue these ideas, the following research issues ought to be investigated:

1. The analysis framework described needs to be developed in more detail; its effectiveness in identifying useful results needs to be evaluated in one or more case studies.
2. The relevance of the research methods described needs to be evaluated with regard to the problems described for large cities in China.
3. Computing resources required to perform such an analysis need to be determined, and their feasibility determined.

7. CONCLUSIONS AND RECOMMENDATIONS

Opportunities clearly exist for China to avoid the errors and shortcomings of traditional academic training and professional practice in North America and Europe, if only because Chinese travel forecasting practice remains at an early stage of development. Opportunities for China's travel forecasting professionals to contribute new approaches, therefore, are unconstrained by past traditions and practices. The challenges of developing new solutions to difficult problems, where others have failed, are quite evident. Associated with this challenge is the need for modelers to acquire the necessary mathematical and scientific expertise. The effort required in this regard is formidable, but quite feasible to accomplish, given the pool of talented Chinese engineers available.

China's professional practitioners need to examine anew the basic problems of urban travel forecasting and urban transportation system design, unbiased by the traditional practices of Western countries. Practitioners should build on what has come before them, but critically assess its applicability and relevance to China's cities. Moreover, given its large and diverse cities, much can be learned from China's urban conditions, which offer new opportunities for experimentation. If successful, China's universities should be able to contribute to urban transportation methods and policy at the international level, as new approaches and insights are needed in all regions of the world, and especially in less developed countries in Asia and Africa.

The following advice is offered to China's emerging travel forecasting expertise. First, critically assess and evaluate all models and software systems before applying them. Second, ask penetrating questions about assumptions and properties of models in order to understand them fully, and to determine their strengths and weaknesses. Third, maintain an open mind about competing software products, since claims may be exaggerated. Fourth, seek opportunities to advance the state of the art of model formulation and solution. You are bound to succeed, and we look forward to your contributions to international travel forecasting practice and the state of the art.

ACKNOWLEDGEMENTS

In June 2006 we visited several transportation planning agencies in China's largest cities, and engaged senior planners in technical discussions concerning their travel forecasting procedures. We also presented our views and understanding of the difficult tasks they face, based on our academic and professional experience in the United States. Following these meetings, we decided to write a paper summarizing what we had learned from these discussions, as well as our perspective on travel forecasting in China. As we had not planned to write this paper initially, we did not inform the participants in these discussions in advance of this intention. For this reason, we do not state here the names of the agencies with whom we met, and we have related our impressions at a relatively general level.

The first author visited China for the first time in 2001. On that occasion he presented several seminars at universities in Beijing and Hong Kong, but did not have an opportunity to interact with professional travel forecasters. His recent second visit was motivated by a desire to inform himself about the level and originality of travel forecasting in large Chinese cities. The second author was born and educated in China, but only entered the urban transportation field when she enrolled in graduate study at the University of Illinois at Chicago in 2000. Following completion of the M.S. degree in civil engineering, she worked professionally in American travel forecasting consulting practice for three years before returning to China in 2005. During a portion of 2005-2006 she was engaged in travel forecasting software application and support activities in China.

The support of the first author's local travel expenses and airfare by the following organizations is gratefully acknowledged: Beijing University of Aeronautics and Astronautics; Beijing Jiaotong University; China Academy of Urban Planning and Design, Tongji University and PTV China. The second author was an employee of PTV China at the time the meetings and discussions reported in this paper were conducted; however, she was on leave during that period. Results reported in Figures 8 and 9 are based on a model implemented in VISUM by Wolfgang Scherr, PTV America.

Section 4 is based on an unpublished paper jointly authored by Hillel Bar-Gera and the first author. We are grateful for his collaboration and comments on this paper. Huw Williams offered comments and encouragement on the preparation of this paper. The advice of two anonymous referees is also gratefully acknowledged. The opinions presented here are solely those of the authors.

REFERENCES

Bar-Gera, H. (2002) Origin-based algorithm for the traffic assignment problem. Transportation Science, 36, 398-417.

Bar-Gera, H. and Boyce, D. (2003) Origin-based algorithms for combined travel forecasting models. Transportation Research Part B, 37, 405-422.

Beckmann, M., McGuire, C.B. and Winsten, C.B. (1956) Studies in the Economics of Transportation. Yale University Press, New Haven, Connecticut.

Boyce, D. (1984) Network models in transportation/land use planning. In M. Florian (ed.), Transportation Planning Models, North Holland Publishing Company, Amsterdam, pp.475-498.
Boyce, D. (2002) Is the sequential travel forecasting procedure counterproductive? Journal of Urban Planning and Development, ASCE, 128, 169-183.
Boyce, D. (2006) An account of a road network design method: Expressway spacing, system configuration and economic evaluation. In X. Feng (ed.), Infrastrukturprobleme bei Bevölkerungsrückgang (Infrastructure Problems under Population Decline), Schriften zur öffentlichen Verwaltung und öffentlichen Wirtschaft, Bd. 202, Berliner Wissenschafts-Verlag, Berlin, forthcoming.
Boyce, D. and Bar-Gera, H. (2003) Validation of multiclass urban travel forecasting models combining origin-destination, mode, and route choices. Journal of Regional Science, 43, 517-540.
Boyce, D. and Bar-Gera, H. (2006) Solving the sequential travel forecasting procedure with feedback. Sixth International Conference of Chinese Transportation Professionals, Dalian, China.
Boyce, D.E. and Cote, R.W. (1966) Verification of land use forecasting models: procedures and data requirements. Highway Research Record, 126.
Boyce, D. and Zhang, Y. (1997) Calibrating a combined model of trip distribution, modal split, and traffic assignment. Transportation Research Record, 1607, 1-5.
Boyce, D.E., Day, N.D. and McDonald, C. (1970) Metropolitan Plan Making. Regional Science Research Institute, Philadelphia.
Boyce, D., Ralevic-Dekic, B. and Bar-Gera, H. (2004) Convergence of traffic assignments: How much is enough? Journal of Transportation Engineering, ASCE, 130, 49-55.
Creighton, R.L. (1970) Urban Transportation Planning. University of Illinois Press, Urbana.
Evans, S.P. (1976) Derivation and analysis of some models for combining trip distribution and assignment. Transportation Research, 10, 37-57.
Fernandez, J.E., de Cea, J and Malbran, H. (2005) Demand responsive urban public transport system design: methodology and application. Working paper, Departmento de Ingeniería de Transporte, Pontifícia Universidad Católica de Chile, Santiago, Chile.
Lam, W.H.K. and Huang, H.J. (1992a) A combined trip distribution and assignment model for multiple user classes. Transportation Research Part B, 26, 275-287.
Lam, W.H.K. and Huang, H.J. (1992b) Calibration of the combined trip distribution and assignment model for multiple user classes. Transportation Research Part B, 26, 289-305.
Lowry, I.S. (1964) A Model of Metropolis. RM-4035-RC, Rand Corporation, Santa Monica, California.
Pickerell, D. (1989) Urban Rail Transit Forecast versus Actual Ridership and Costs. Urban Mass Transportation Administration, US Department of Transportation, Washington, DC.
Powell, W.B. and Sheffi, Y. (1982) The convergence of equilibrium algorithms with predetermined step sizes. Transportation Science, 16, 45-55.
Sacks, J., Welch, W., Mitchell, T. and Wynn, H. (1989) Design and analysis of computer experiments. Statistical Science, 4, 409-435.
Siegel, J.D., De Cea, J., Fernández, J.E., Rodríquez, E.E. and Boyce, D. (2006) Comparisons of urban travel forecasts prepared with the sequential procedure and a combined model. Networks and Spatial Economics, 6, 135-148.

Wardrop, J.G. (1952) Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers, Part II, 1, 325-378.
Williams, H.C.W.L. (1977) On the formation of travel demand models and economic evaluation measures of user benefit. Environment and Planning A, 9, 285-344.
Wu, J.H., Florian, M. and He, S. (2006) An algorithm for multi-class network equilibrium problem in PCE of trucks: Application to the SCAG travel demand model. Transportmetrica, 2, 1-9.