New Results on Solving the Sequential Travel Forecasting Procedure with Feedback[1]

David Boyce (corresponding author)
Department of Civil and Environmental Engineering
Northwestern University, Evanston, IL 60208
(mailing address: 2149 Grey Avenue, Evanston, IL 60201)
Phone: 847-570-9501; Fax: 775-587-2308
Email: d-boyce@northwestern.edu

Christopher R. O'Neill
Capital District Transportation Committee
One Park Place
Albany, NY, 12205
Phone: 518-458-2161; Fax: 518-459-2155
Email: co'neill@cdtcmpo.org

Wolfgang Scherr
PTV America, Inc.
1300 N. Market Street, Suite 704
Wilmington, DE 19801
Phone: 302-654-4384; Fax: 302-691-4740
Email: wscherr@ptvamerica.com

July 17, 2007

For consideration for presentation and publication by
the Committee on Transportation Planning Applications (ADB50)

Word Count: 4,097 + 8 figures and 1 table

[1] Presented at the 11th Transportation Planning Applications Conference, May 6-10, 2007, Daytona Beach, FL.

Abstract

Travel forecasters generally understand an iterative solution of the Sequential Travel Forecasting Procedure is required to bring specific model inputs and outputs into consistent agreement. In particular, the congested interzonal travel time inputs to the trip distribution and mode choice steps should equal the user-equilibrium travel times obtained from the assignment step. The process of achieving consistency is called "solving the Sequential Procedure with Feedback."

The Capital District Transportation Committee, Albany, NY, maintains a travel forecasting model, which was recently updated to 1000 traffic analysis zones. This model was used to evaluate feedback procedures for three cases drawn from its planning activities. Three alternative feedback solution procedures were applied to the model: a) naïve or direct feedback (no averaging of trip matrices or link flows); b) averaging of trip matrices with Constant Weights; c) Method of Successive Averages (MSA) applied to trip matrices. The convergence of the feedback procedures was measured as follows: Total Misplaced Flow (trip matrix); Relative Gap (route-based user-equilibrium traffic assignment); Root Squared Error (travel cost matrices).

The test results showed that: a) averaging of trip matrices using Constant Weights converges to a single, stable solution with consistent travel costs; b) a single pair of weights is most effective for all three cases; c) neither Naïve Feedback nor MSA is as effective as using Constant Weights; d) the Relative Gaps of the Traffic Assignment reach values less than 1.E-5%.

Tests with different models and software systems are needed to generalize these findings.

INTRODUCTION

Travel forecasters have understood since the first conceptualization of the Sequential Travel Forecasting Procedure (also called the Four-Step Procedure) during the 1950s that an iterative approach is required to bring the inputs and outputs of specific models into consistent agreement with each other (*1*). In particular, the congested interzonal travel costs (shortest route travel times, or skims) which are input to the trip distribution and mode choice steps should equal the user-equilibrium travel costs resulting from the solution of the assignment step. The iterative process of achieving consistency between these input and output travel costs is generally referred to as "solving the Sequential Procedure with Feedback."

Although the need to solve the four-step procedure iteratively was apparent, how to solve this problem has confounded numerous practitioners and researchers. At the 1993 Transportation Planning Applications Conference, the first author organized a session on this topic, which led to the computational experiments reported by Boyce et al. (*2*). Subsequently, Comsis (*3*) undertook a study of how to incorporate feedback in travel forecasting; the results were rather inconclusive. Recently, however, Boyce, in collaboration with Scherr, performed computational experiments that appeared to be more promising (*4*, *5*).

In all of these investigations, the apparent difficulties stemmed from the question of just how to perform the feedback calculations in the Sequential Procedure. The experiments of Boyce et al. (*2*) and Comsis (*3*) demonstrated that it is ineffective to feed back the travel costs directly from the previous loop (Naïve Feedback, also termed as Direct Feedback). Evidently, some sort of averaging between successive loops is required. But what should be averaged? Various investigators have averaged link flows, link cost, or even link speeds.

To gain a clearer view of the problem, let us reconsider what we are trying to achieve. We wish to find a multimodal trip matrix that depends in part on interzonal generalized modal travel costs (linearly weighted sums of travel times, operating costs, tolls or fares). For congested modes (typically, car), we want these travel costs to be either costs of the shortest routes, or the average costs of used routes, depending upon whether a link-based or a route-based assignment algorithm is applied. To achieve our objective, we seek a multimodal trip matrix, partly dependent on modal travel costs, which if assigned to the multimodal network, yields those same costs. This problem statement suggests that we ought to focus on finding the multimodal trip matrix that satisfies this criterion, rather than focusing on link flows or costs. This line of thinking led us to define a feedback procedure that involves averaging of trip matrices. The final matrix should, when assigned to the network, yield a matrix of shortest travel costs that implies the same trip matrix. Persons with a mathematical background may recognize this as the statement of a fixed point problem; see Bar-Gera and Boyce (*6*, *7*).

In this paper we report on tests of three ways of averaging successive trip matrices:

1. averaging with Constant Weights (sometimes referred to as fixed weights);
2. averaging with the Method of Successive Averages (MSA), in which the weight on each new solution matrix decreases with each feedback loop.
3. Naïve Feedback (Constant Weight averaging with a weight of 1.0 on the new matrix).

The paper is organized as follows:

1. description of the model and the cases used in these tests;
2. description of the feedback procedure with an accompanying flowchart;
3. findings regarding the convergence of the trip matrix to a unique solution;
4. convergence of zone-to-zone travel costs and the traffic assignments;
5. effect of the initial cost assumption and the level of congestion on convergence;

6. conclusions, recommendations and a call for further studies by practitioners.

DESCRIPTION OF THE MODEL

The Capital District Transportation Committee (CDTC) is the MPO for the New York State counties of Albany, Rensselaer, Saratoga and Schenectady, which presently have a total population of about 800,000, primarily in the central cities of Albany, Troy, Schenectady and Saratoga Springs, and the contiguous suburban development, as shown in Figure 1. CDTC has 25 members, including the New York State DOT, New York State Thruway Authority, Capital District Transportation Authority, Capital District Regional Planning Commission, Port of Albany, Albany International Airport and elected officials from counties, cities and towns.

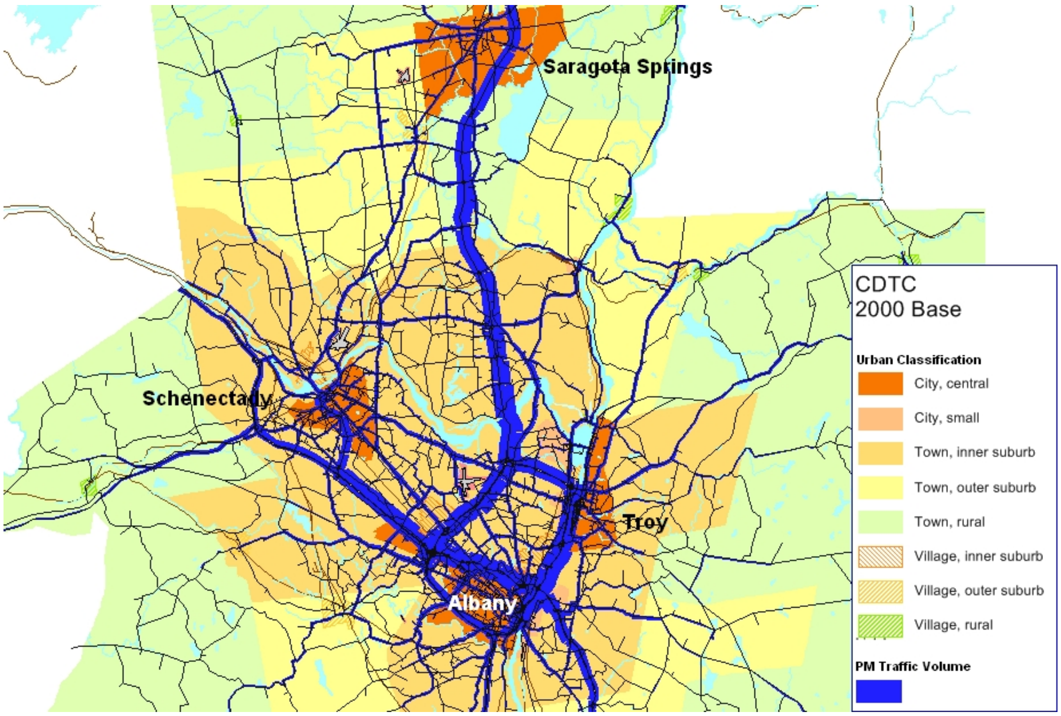

FIGURE 1 CDTC Travel Forecasting Model – Network and Urban Classification.

CDTC made extensive use of its regional travel forecasting model in preparing its New Visions Regional Transportation Plan, as well as for its Community and Transportation Planning Linkage Program, evaluation of Transportation Improvement Program project candidates, and development of performance measures. CDTC also provides traffic forecasts to New York State DOT for project development. Forecasts assume steady progress in the integration of land use and transportation plans, urban reinvestment, demand management and pedestrian, bicycle and transit access.

The current travel model includes five trip purposes through trip generation and trip distribution, which are augmented by through traffic. The trip distribution models are doubly-constrained gravity models with negative power functions. The only mode currently represented in the model is auto travel; public transit is not yet part of the application. Auto assignment is computed for the PM peak-hour utilizing capacities defined on Level of Service C volumes, not only on links but also on turns to emulate intersection delay. The network has 1,000 zones, 10,000 links, 4,000 nodes, and 21,000 turns with volume-delay functions. CDTC's modeling practice is based on VISUM 9.5 (*8*); the tests reported here were computed with VISUM 10.0 beta. The computer used for solving the model was an off-the-shelf Windows PC with a 2.0 GHz processor and 2.0 GB RAM memory purchased in 2006.

Three cases are considered in this paper:

1. Base2000: model calibration based on the 2000 Census;
2. Plan2030: the 2030 forecast for the planned network;
3. Base2000x1.5: derived from Base2000 by multiplying the productions and attractions by a factor of 1.5 to obtain a more congested case for these feedback tests.

THE FEEDBACK METHOD

The proposed feedback procedure is depicted in Figure 2 for a problem with a single congested mode (car) operating on a road network. A second mode (transit) operating on a network with fixed travel times and fares is easily added, but omitted here in the interest of simplicity.

The initial solution (Loop 1) to the four-step procedure is represented by Box #2, consisting of trip distribution and assignment, given an initially assumed travel cost matrix and the zonal origin and destination totals by trip purpose in Box #1. Box #3 represents the general solution of the trip distribution model in Loop k based on the travel cost matrix from the last assignment in Loop k-1. In Box #4 this new trip matrix from Box #3 is averaged with the trip matrix from the previous Loop k-1, yielding the averaged trip matrix of Loop k.

The averaging can be performed with Constant Weights (CW), with a weight of w on the averaged matrix and a weight of $(1-w)$ on the new matrix. Or, the Method of Successive Averages (MSA) can be applied; using this terminology the weight on the averaged matrix is $w=(k-1)/k,$ and the weight on the new matrix is $(1-w)=1/k,$ which is sometimes called the step size. Naïve Feedback is a special case of Constant Weights with $w=0$ and $(1-w)=1.$ The averaged matrix is then assigned to the road network (Box #5), and a convergence test is performed in Box #6 to test whether the two matrices averaged in Box #4 are effectively equal. MSA can be mathematically proven to converge to the desired consistent solution, but convergence may be slow. Averaging with Constant Weights has not been proven to converge.

Two useful convergence measures, Total Misplaced Flow (TMF) and Root Squared Error (RSE), are defined in the flowchart. (As both measures yield essentially the same pattern, only TMF is shown in the results presented below.) If the convergence measure selected is not sufficiently small, the procedure returns to Box #3, and begins again. Note that only trip matrices are averaged; link flows, link costs and zone-to-zone costs are computed from the assignment in Box #5, and used directly in solving the trip distribution model again.

COMPUTATIONAL RESULTS

Convergence of the Solution Procedure

For each of the three cases, solutions with various feedback weights were computed: Constant Weights with values of w ranging from 0.0 to 0.7, and MSA. TMF in log scale is plotted versus computational time for the Base2000 case in Figure 3; similar results were obtained for the two other cases. Figure 4 shows a comparison of the Base2000 case with the Plan2030 and Base2000x1.5 cases for the most effective Constant Weights (w = 0.25) as well as MSA and Naïve or Direct Feedback.

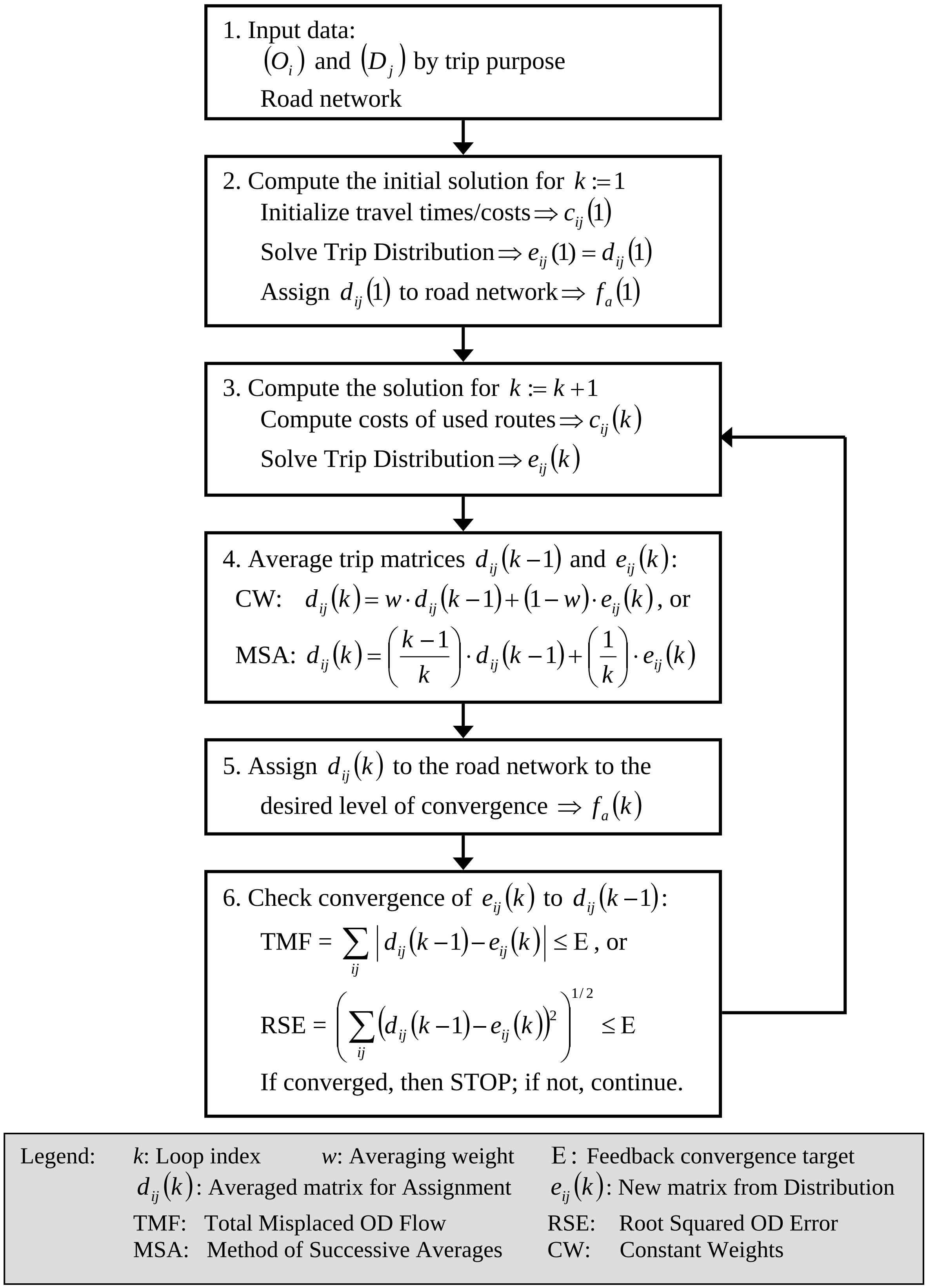

FIGURE 2. Feedback Procedure for Averaging Origin-Destination Matrices.

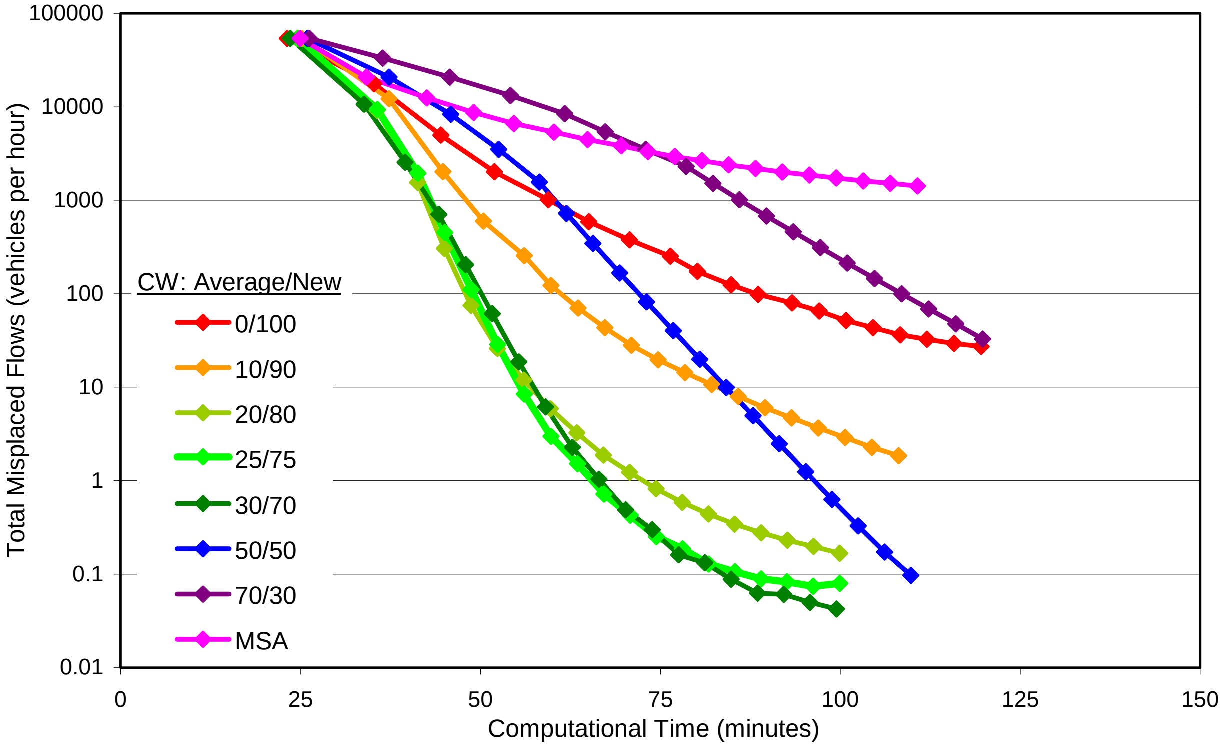

FIGURE 3 Convergence of Trip Matrices for Base2000.

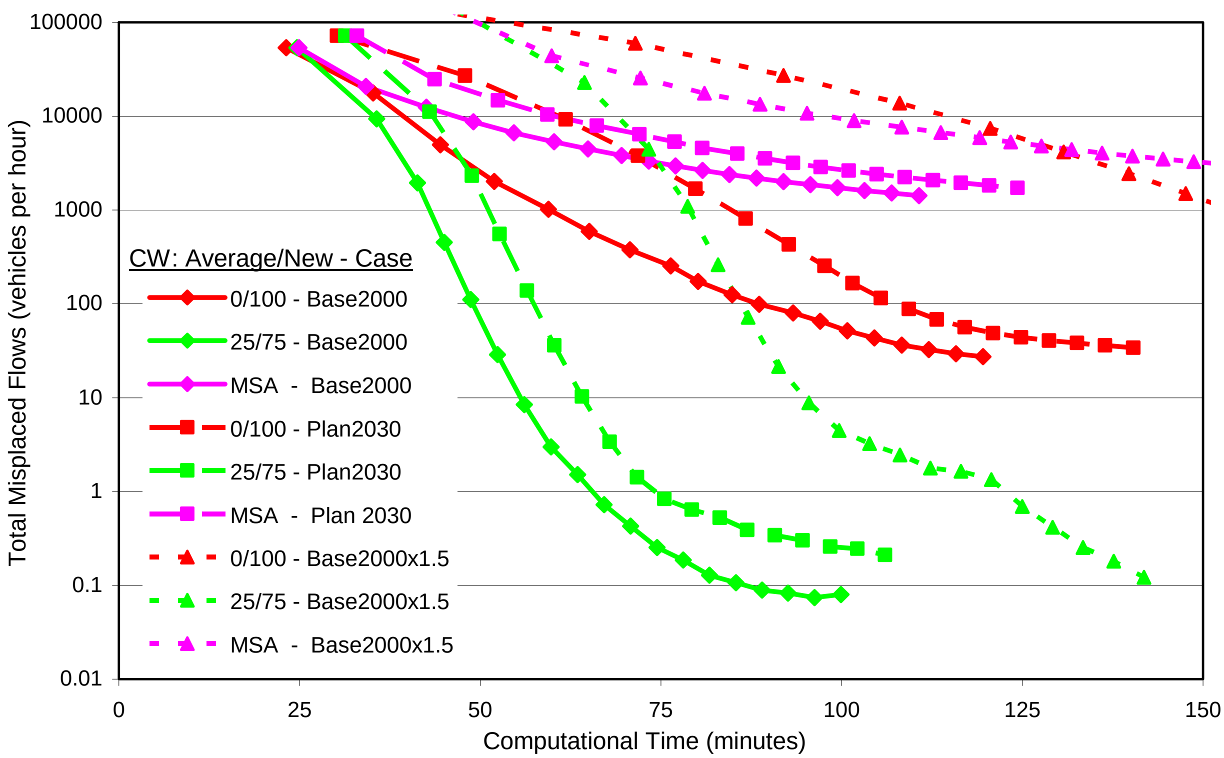

FIGURE 4 Convergence of Trip Matrices for All Three Cases.

The patterns in the reduction of TMF are similar for the three cases for values of *w* of 0.2 to 0.5. Values of *w* of 0.0 and 0.1 are much less effective, especially in the most congested case, Base 2000x1.5. The MSA method is ineffective in all three cases, as are values of 0.7 and higher. Notice that a *w* value of 0.5 is quite effective at higher solution times, but somewhat ineffective earlier. In addition to the stability of the overall patterns, an important finding is that a *w* value of 0.25 is best for all three cases. This result implies that the same Constant Weights might be applied to all future plan scenarios; adjustment of the Constant Weights for each scenario would be unnecessary.

To consider how many feedback loops are required for practical applications of the method, see Figure 5. For any method, the minimum number of feedback loops is three: the initial solution; a second loop to determine the level of disparity from the initial solution; and a third loop to check whether averaging of the first two solutions is adequate. The real question, then, is how many more loops are needed. For a *w* value of 0.25, at Loop 4 TMF is reduced to 4,500 vph, or 1.4% of the total origin-destination (OD) flow of 313,310. Loop 5 reduces TMF to 1,100 vph or 0.4%, which seems sufficient for comparisons of alternative plans. In contrast, MSA results in a TMF of over 17,600 vph or 6% of total flow after five loops with a similar solution time for a *w* value of 0.25. Naïve Feedback (*w* = 0.0) has a TMF of 27,200 vph, or 9% of total flow with a higher solution time. The reason that Naïve Feedback is much slower is that more iterations are required to reach the specified level of assignment convergence. In conclusion, we recommend TMF should be less than 1% of total OD flow.

For all three cases, five feedback loops with a *w* value of 0.25 are sufficient to achieve a reduction in TMF to less than 1% of total flow. At this point in the solution process, however, the rate of improvement in TMF from additional feedback loops is very high, so that one or two more loops may well be worthwhile for this best value of *w*. In contrast, the rate of improvement in TMF is relatively low for MSA and Naïve Feedback.

Unique Trip Matrix Solution

An important question raised about these results concerns whether the proposed averaging procedure converges to the same trip matrix for the various weights applied. To investigate this matter, the feedback procedure was solved with a *w* value of 0.5 for 100 loops, yielding a precisely converged solution with TMF = 0.0025. Then, the procedure was applied with values of *w* ranging from 0.0 to 0.7, as well as MSA for 20 loops. The sum of the absolute differences in cell values between each trip matrix and the precisely converged matrix was computed at the conclusion of each loop.

Figure 6 shows the results of this computational experiment. For all of the Constant Weights except *w* values of 0.0 and 0.7, the trip matrix converges to a very close approximation of the precisely converged matrix. For the Naïve Feedback and MSA solutions, convergence is not achieved; however, it might be achieved after a very large number of feedback loops. In Figure 6 the x-axis shows the number of feedback loops, rather than computational time, because the computational times are not comparable to Figures 3-5. Moreover, plotting the result in this manner facilitates the comparison of the Total Differences in Flow at each feedback loop.

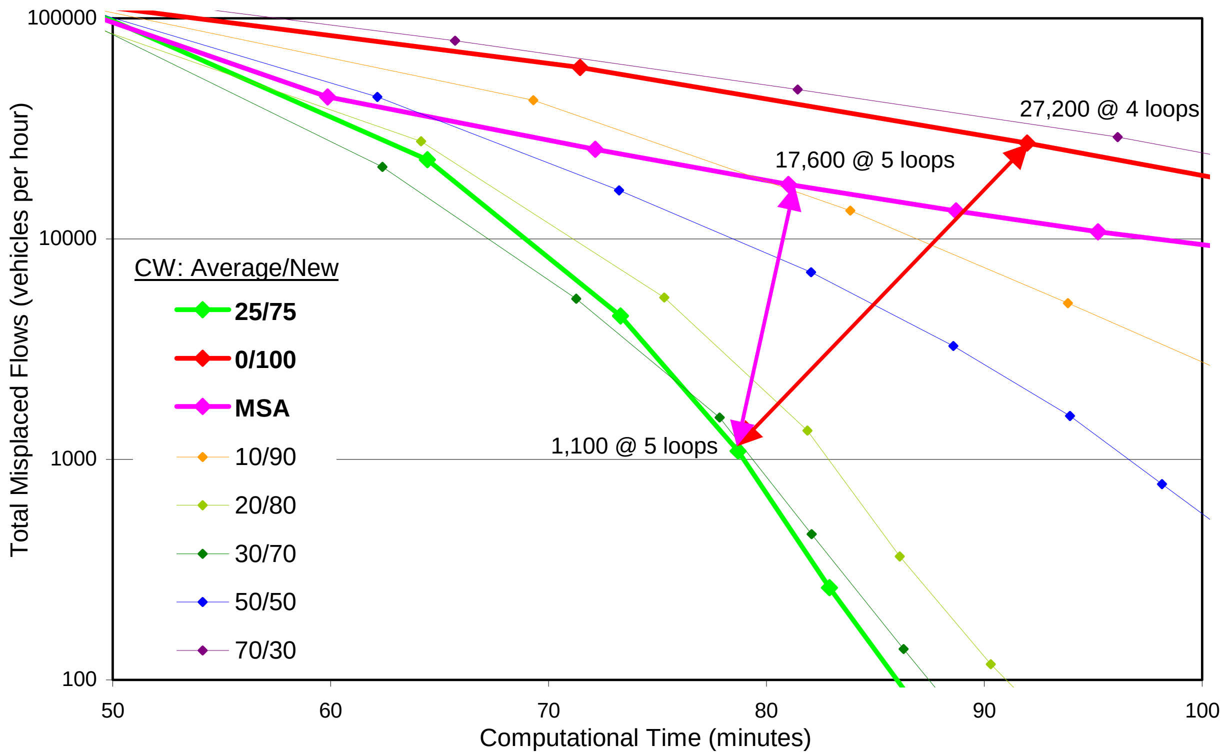

FIGURE 5 Recommended Stopping Criterion for Trip Matrices for Base2000x1.5.

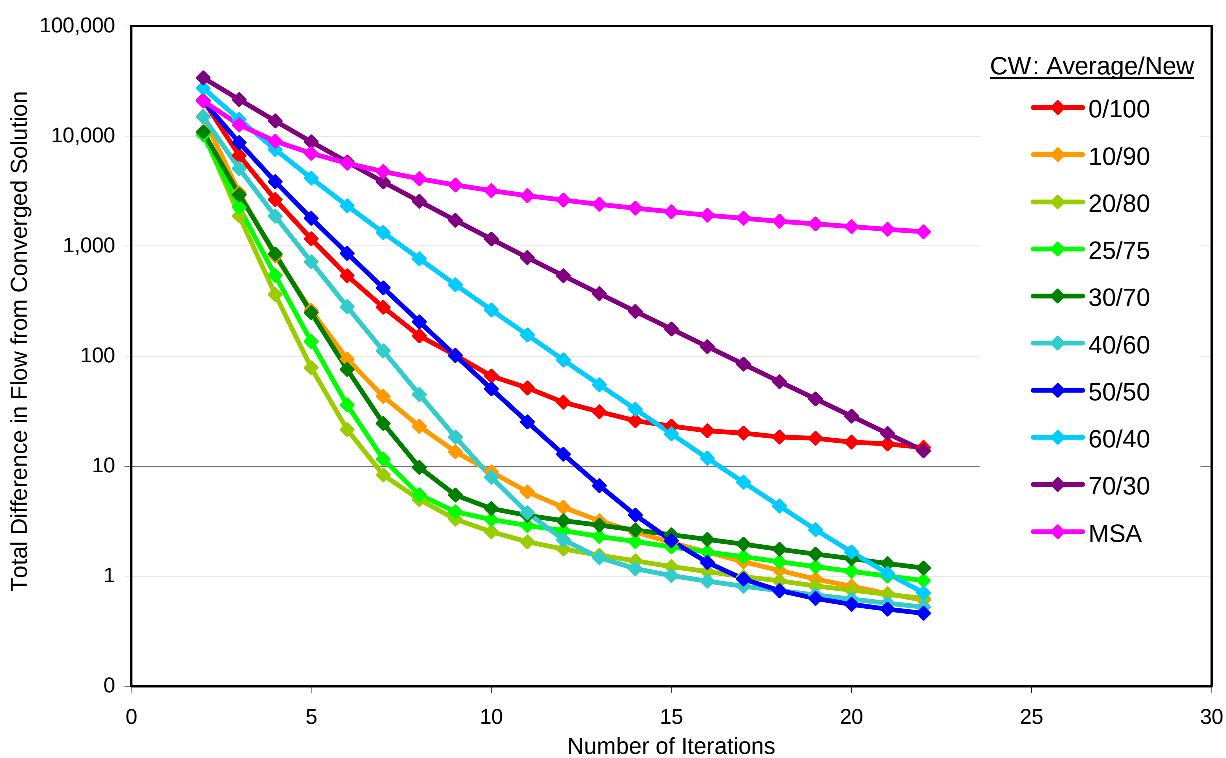

FIGURE 6 Comparison of Ten Solutions with the Converged Solution for Plan2030.

Convergence of Travel Cost Matrices

In addition to examining the convergence of trip matrices, which partly depend on the zone-to-zone travel costs, we should also determine whether the successive travel cost matrices are converging to a stable value. To examine whether this criterion is satisfied, we computed the root squared error (RSE) of pairs of successive travel cost matrices. The results for Plan2030 are shown as Figure 7. Similar plots were obtained for the other two cases. Note that the cost matrices computed from the route-based assignment are the flow-weighted *average* costs over all used routes, not the costs of the *current shortest routes* as found in link-based assignments.

In the initial solution, free flow travel times are used to compute the first trip distribution. The second trip matrix is then based on the average travel costs from a multi-path assignment of the initial trip matrix. The RSE for these two travel cost matrices is 17,300 minutes, as shown by the first cluster of points in the upper left of Figure 7. For the best *w* value of 0.25, RSE decreases sharply to a value less than 1. For MSA and Naïve Feedback, RSE decreases more slowly, again suggesting these methods do not converge as quickly as the best Constant Weights.

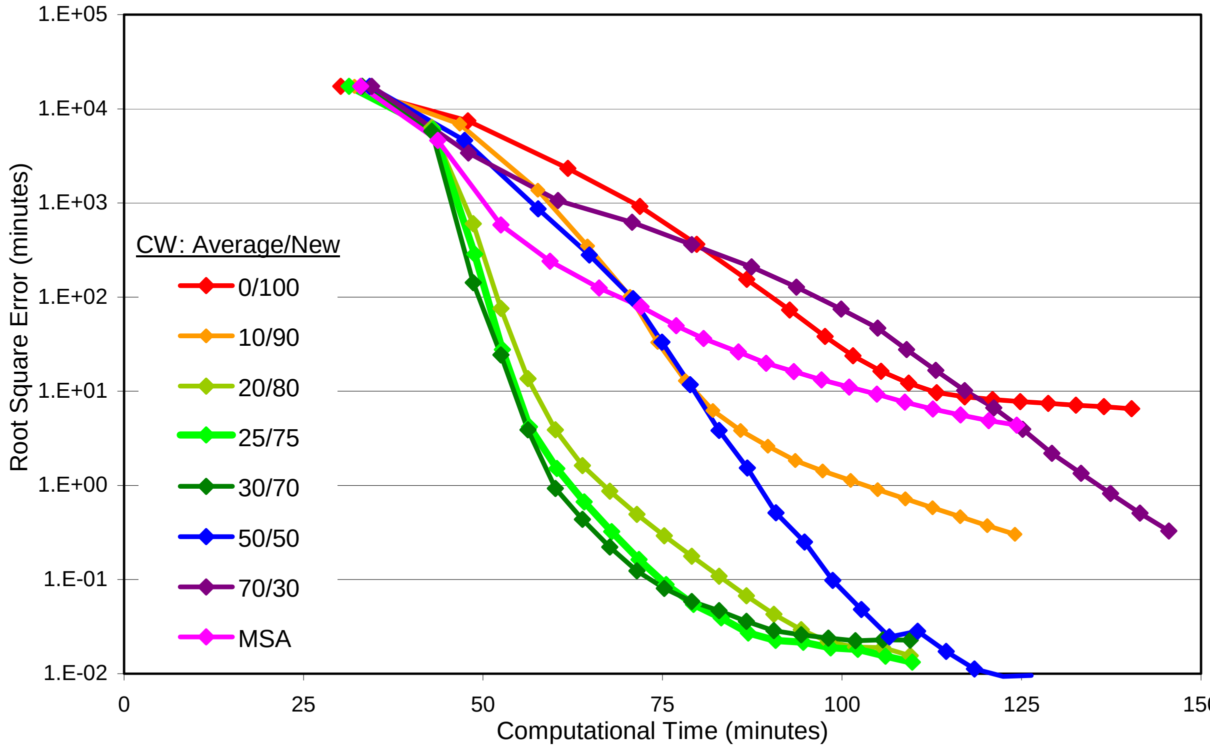

FIGURE 7 Convergence of Travel Costs for Plan2030.

Convergence of Traffic Assignments

Each loop of the iterative Sequential Procedure requires the solution of the Traffic Assignment problem, which is an iterative procedure in itself. In these tests the assignment is solved with a route-based user-equilibrium (UE) method, as described by Bothner and Lutter (*9*) and Schittenheim (*10*), rather than the link-based methods generally found in other travel forecasting software systems. This method has been provided in VISUM since the 1980s.

The route-based algorithm, in contrast to link-based methods, stores all routes that belong to the current solution. The knowledge of routes, and of the relationships between routes and links, is used to shift OD flows among alternative routes until the route costs are equal and minimal for each OD pair; this shifting of route flows is called "balancing." In its outer iteration the assignment algorithm performs: 1) a shortest route search (for all OD pairs); 2) several inner iterations of balancing over all OD pairs, including simultaneously updating the link and turn travel times; and finally 3) a convergence test defined on the maximum Relative Gap (RG). The user of the route-based assignment controls the maximum number of outer loops, the maximum number of balancing iterations per main loop and the level of RG as a stopping criterion. Our settings for these tests are: 1) number of outer loops – unlimited; 2) maximum number of balancing iterations – 3; and 3) RG < 1.E-5 (for the initial solution, RG < 1.E-1).

The convergence of the assignment steps is shown in Figure 8. The level of assignment convergence (measured by Relative Gap) achieved is far better than what most practitioners are able to achieve with link-based methods (typically on the order of 1.E-3). For the most efficient feedback methods (CW with *w* values between 0.1 and 0.4), the Relative Gap decreases to 1.E-7 after 15 to 20 feedback loops in all three cases. After six feedback loops a very stable assignment with a RG less than 1.E-6 is reached, which means that extremely little "noise" would be found if two plan alternatives were compared.

Note in the later feedback loops the Relative Gap decreases well beyond the stopping criterion of 1.E-5. During each feedback loop, the route flow from the previous assignment is adjusted with the updated averaged trip matrix (called a "reload"). After each reload the algorithm performs a minimum of two route searches and a several balancing steps, which is the reason why the assignments continue to improve below the stopping criterion.

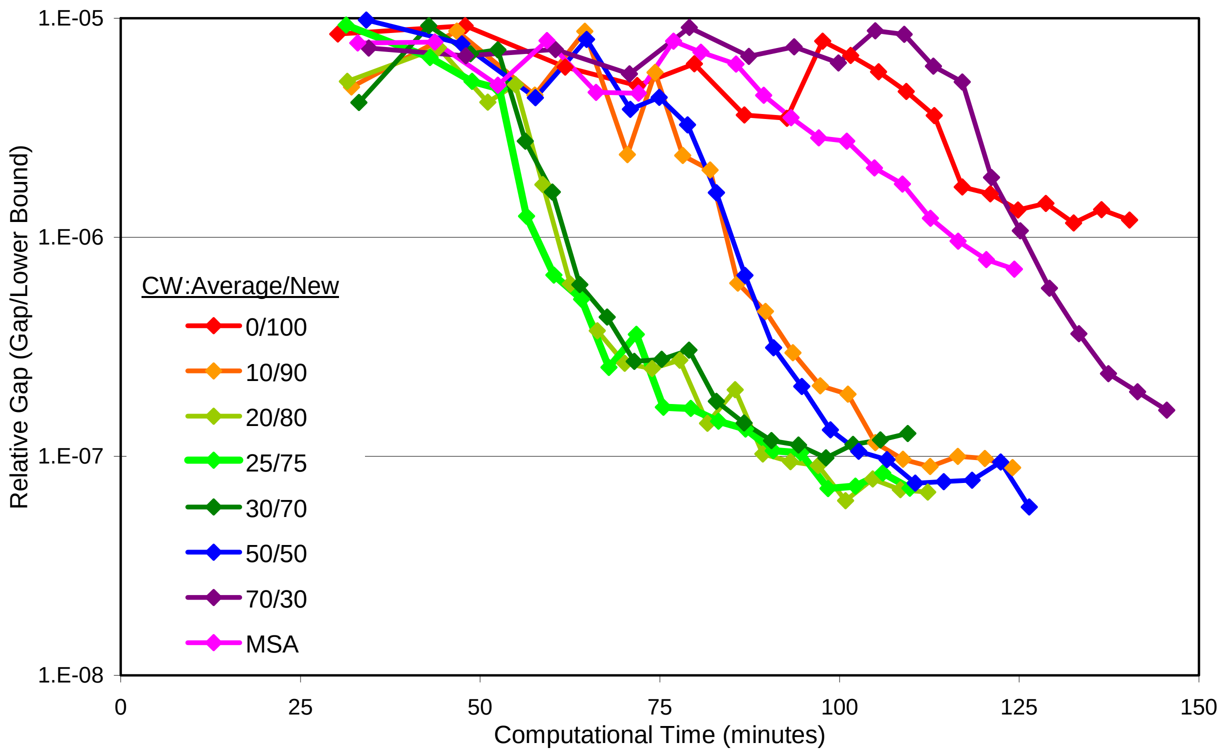

FIGURE 8 Convergence of Relative Gaps for Plan2030.

Effect of Initial Travel Cost Assumption on Convergence
In all of the results reported above, free flow travel costs were used to find the initial trip matrix. Suppose a travel cost matrix from a related problem was used instead. Would the convergence of the feedback procedure be improved? To explore this question, the feedback procedure was initiated with a previous solution to a similar problem. For the best Constant Weights, relatively little improvement was observed with this approach.

Effects of the Level of Congestion on Convergence
The case studies for this paper were performed on a zone system and road network for a medium -sized metropolitan region (800,000 population) with moderate levels of congestion during the peak hour in the base year. Practitioners from larger metropolitan areas have questioned whether the findings are applicable to more congested conditions. Although unable to answer this question, we wish to document the congestion levels for the case studies for comparison with future findings. Table 1 shows some key indicators of the size of the problem, the number of links and turns with volumes in excess of 1.25, and the space-mean-speed (VMT/VHT). The capacities in the volume-delay functions used in this model are set for Level of Service C. Therefore, a volume greater than or equal to 1.25 times the capacity is equivalent to Level of Service E or F.

In the CDTC model, travel times are not only based on link delays, but also on intersection delays, where intersection delays are modeled by turn-based volume-delay functions. Table 1 shows that turn delays contribute more to overall network delay than link delays. Our experience with other link-and-turn based models has shown that during the equilibrium assignment and feedback iterations the v/c ratios are higher on the turns than on the links.

TABLE 1 Characteristics of the CDTC Network and Case Study Solutions

Indicator variable	**Base2000**	**Plan2030**	**Base2000x1.5**
Total origin-destination flow (veh/hr)	207,540	233,100	311,310
No. of links with capacities	9,760	9,790	9,760
No. of links with v/c > 1.0	280	420	830
No. of links with v/c > 1.25	80	140	340
No. of turns with capacities	21,300	21,390	21,300
No. of turns with v/c > 1.0	250	350	660
No. of turns with v/c > 1.25	100	300	150
Space-mean-speed (mph)	34.1	31.9	27.4

CONCLUSIONS, RECOMMENDATIONS AND FURTHER STUDIES
From the tests conducted for three cases with the CDTC model, we draw the following conclusions:

1. Averaging the trip matrix using Constant Weight values of *w* in the range of 0.2 to 0.5 yields stable and highly converged solutions to the problem of solving the Sequential Procedure with Feedback. This result agrees with the findings of Bar-Gera and Boyce (*7*) for a research model with well-defined convergence properties.
2. The same *w* values of 0.25 were highly satisfactory for three cases with quite different congestion levels. This finding is significant because it suggests weights that are good for one case can be transferred to another case without extensive testing.

3. Performing feedback without averaging (Naïve Feedback) is relatively ineffective and should not be used. The Method of Successive Averages is much less effective than using Constant Weights in these tests. MSA should only be used if Constant Weights have been shown to be ineffective, as could happen with another model.
4. For the tests conducted, performing five feedback loops was effective in reaching convergence, as measured by Total Misplaced Flow. Additional loops improved the convergence to some extent; divergence of the solution was not observed.
5. Tests are required for each practitioner's model to determine the effective number of feedback loops. As a general guideline, we suggest that the ratio of Total Misplaced Flow to the total flow among all origins and destinations should be less than 1%.
6. Procedures using link-based algorithms for auto assignment may need more feedback loops because the assignment is often unable to achieve precise convergence levels.

The experience accumulated to date using the VISUM software system pertains to three cases solved with the CDTC model. Additional tests with more complex models and other software systems are needed to generalize these findings further. Practitioners are urged to perform their own tests and report them in a manner such that findings across models, networks and software systems can be compared.

Acknowledgement We thank Hillel Bar-Gera for advice on the design of this research, and for comments on the findings. Comments of several practitioners are also gratefully acknowledged.

REFERENCES

1. Carroll, Jr., J. D., and H. W. Bevis. Predicting Local Travel in Urban Regions. *Papers and Proceedings, The Regional Science Association*, Vol. 3, 1957, pp. 183-197.
2. Boyce, D., M. Lupa, and Y. Zhang. Introducing 'Feedback' into the Four-step Travel Forecasting Procedure vs. the Equilibrium Solution of a Combined Model. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1443*, TRB, National Research Council, Washington, D.C., 1994, pp. 65-74.
3. Comsis Corporation. *Incorporating Feedback in Travel Forecasting: Methods, Pitfalls and Common Concerns*. Publ. DOT-T-96-14. FHWA, U.S. Department of Transportation, 1996.
4. Bar-Gera, H., and D. Boyce. Solving the Sequential Procedure with Feedback. Presented at the 6th International Conference of Chinese Transportation Professionals, Dalian, China, 2006.
5. Boyce, D., and C. Xiong. Forecasting Travel for Very Large Cities: Challenges and Opportunities for China. *Transportmetrica*, Vol. 3, 2007, pp. 1-19.
6. Bar-Gera, H., and D. Boyce. Origin-based Algorithms for Combined Travel Forecasting Models. *Transportation Research*, Vol. 37B, 2003, pp. 405-422.
7. Bar-Gera, H., and D. Boyce. Solving a Nonconvex Combined Travel Forecasting Model by the Method of Successive Averages with Constant Step Sizes. *Transportation Research*, Vol. 40B, 2006, pp. 351-367.
8. *VISUM 9.5 Manual*. PTV AG, Karlsruhe, Germany, 2007.
9. Bothner, P., and W. Lutter. *Ein direktes Verfahren zur Verkehrsumlegung nach dem ersten Prinzip von Wardrop*. Universitaet Bremen, FB Verkehrssysteme. Arbeitsbericht 1, Bremen, Germany, 1982.
10. Schittenhelm, H. On the Integration of an Effective Assignment Algorithm with Path and Path-flow Management in a Combined Trip Distribution and Traffic Assignment Algorithm. Presented at 18th PTRC Summer Annual Meeting, Sussex, UK, 1990.