Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 10, October 2021:4252-4268

A Review on Mathematical Programming Formulations for Transportation and Land Use Models

T. Pavan Kumar,
Research Scholar,Department of civil engineering, JNTUH
Pavanphd522@gmail.com

Dr. K. M. Lakshmana Rao,
Professor and Director of BICS, Department of civil engineering, JNTUH
Kmlraoin@jntuh.ac.in

Mathematical programming formulations will most likely produce the next generation of transportation, location, and land use models. Trip asslgnment and population placement are both referred to as op t1m1zat lori problems in mathematlca and programming formulatlons in the examples given. First, a trip assignment model with constant llnh costs is shown, and then the same model is changed to highlight the flow-dependent link cost formulatlon's ramifications. For the same reason, a simple, linear model of population distribution has been turned into a complex, nonlinear model that takes into account the varying tastes and perceptions among people who live in a certain area. A nonlinear programmlngformulaton that solves both the trip assignment and the location issue at the same time is then shown. The theoretical merits and practical disadvantages of this approach are briefly mentioned in the paper's concluding section. These strategies may be used in applied planning systems if they can be used to solve practical issues.

There has been considerable refinement of practical methods of forecasting urban location and transportation patterns during the past 10 to 15 years. Although there is continuing discussion and development, and even the best of forecasts are far from perfect, there appears to be a greater consensus on what methods are clearly outmoded and in what directions future efforts should move. This author's views on general progress in the field have already been published *(1).* Among the most sophisticated practical methods of transportation and land use forecasting are the extended spatial interaction models, es- pecially when they are included in comprehensive integrated model systems [see paper by Bly and Webster in this Record and Putman *(2)].*

In addition to these practical developments there have also been important theoretical developments. On the transportation side these include the development of discrete choice models, especially for travel demand and mode choice *(3),* and the shown the existence of clear errors in prior practice; and othersmay offer substantial improvements for future applications.Past experience suggests that there is a lag of 10 years,sometimes more, between the initial development and subse- quent practical application of new techniques in transportation and land use forecasting. Thus, although there have been some attempted applications of these methods *(7 ,8),* they are farfrom being the accepted norm. The purpose of this paper is topresent some illustrations of the mathematical programmingformulations along with some simple numerical examples. Theintent is to show some of the benefits, both practical andtheoretical, of these formulations and to provide the practical planner with an introduction to this promising new area.

NETWORKS AND TRIP ASSIGNMENT

In thediscussionthatfollows,extensiveusewillbemadeofthe data describing Archerville, a simple five-zone numerical ex- ample. Table 1 give the land use and socioeconomic data for Archerville. Figure 1 shows the Archerville highway system.

Shortest Path Problem

A frequently encountered problem in transportation and loca- tionanalyses is that of finding the shortest path from one node to another over the links of a network. This is a problem that can be considered as a linear programming problem. The equa- tion form is (1)
subject to

JJ, pigs $X_{ki} = \Big\{$ $\textbf{Min: } Z = \sum_i \sum_j$

$c_{ij}\ X_{ij}$

development of mathematical programming formulations of the traffic assignment problem *(4).* On the location side the development of utility theory as a basis for location models *(5)*and the general discussion of mathematical programming mod- elsas alternate or underlying structures for spatialinteraction—Z -1 if i =destination $\sum_j X_{ki}$
k 0 otherwise

$X_{ij} \geq 0 \quad (\forall\ i, j)$ (2)
(3)

models *(6)*were major developments. Some of these develop- mentsare important principally because they provide an im- proved underpinning of existing practical methods; some have wherec_{ij}iscostoftraversinglink*i*,yandX;jisflow(trips)on
link *i*, y.

TABLE 1 ARCHERVILLE—LAND USE AND SOGOECONOMIC DATA

	LandUseData					SocioeconomicData					
Zone	Resi-dential	Commer-cial	Indus-trial	Vacant	Total	Cmnmercial Employment	Industrial Employment	Total Employment ∞	M House-holds	HI House-holds	Total Population
1	2.5	1.0	1.0	0.5	5.0	150	150		200	100	860
2	3.0	1.0	1.5	0.7	6.2	200	150	350	300	50	1,050
3	1.0	2.0	3.0	0.8	6.8	100	400	∞	150	50	585
4	2.5	1.0	0.0	4.5	8.0	200	0		100	300	550
5	1.5	0.5	0.0	6.6	8.6	100	0	I£D	50	150	515
Total						750	700	1,450	800	650	3,560

Employee-HouseholdCrossTabulation

	LI	HI	Total
Commercial	400	350	750
Indusuial	400	3%	700
	800	650	1,450

Employee-Household ConversionMatrix

	LI	HI	Total
Commercial	.533	.467	1.00
Industrial	.571	.429	1.00

None: M = low-income and HI = high-income.

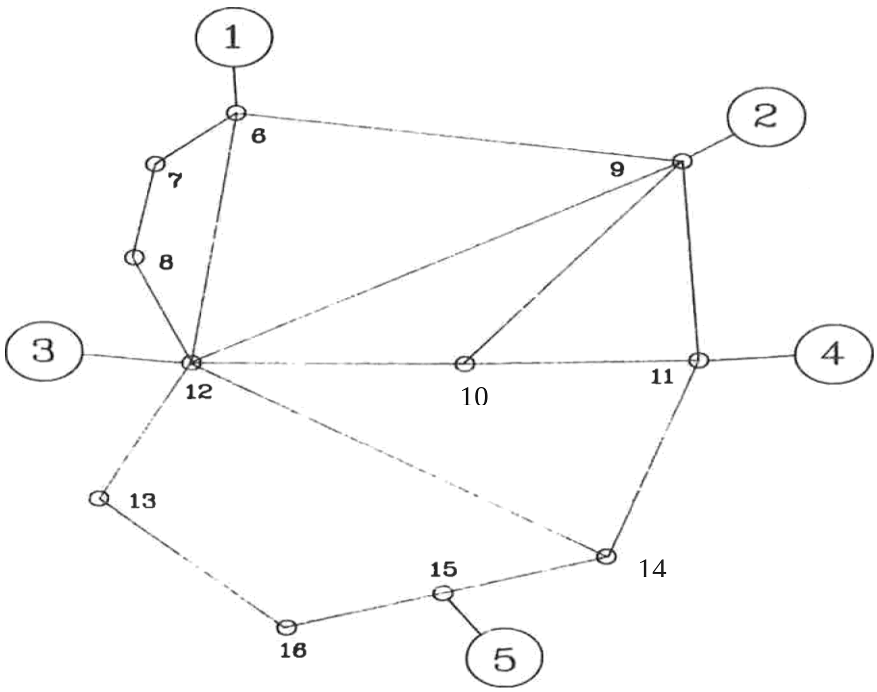

FIGURE 1 Archervillehighway system.

network node, balance. Thus for each node the total flow of trips into the node minus the total flow of trips out of the node must equal the net trips supplied (or demanded) at the node. The last constraint equation (Equation 3) is simply the non- negativity requirements that prohibit negative flows. It can readily be seen that writing the shortest path problem in this form yields a rather good sized problem. In the objective function the number of terms equals the number of links in the network. There must be a constraint equation (Equation 2) for each network node and a constraint equation (Equation 3) for each network link.

For practical applications there are many fast algorithms for solving this problem. However, it is worth noting here that, because this is just a simple linear program, it can be solved by the well-known simplex method. If it were done that way, it would be necessary to solve the problem once for each origin- destinationpath thatwaswanted.Furthernotethatthisproblem only addresses the situation for fixed link costs and flow vol- umes, which must, of course, be known in advance of any attempt to solve for the minimum path orpaths.

Minimum-Cost **Flow Problem**

Another type of linear programming problem is that known as the minimum-cost flow problem. The Archerville data may be used as an example. Assume that there are a known number of households of each income group in each zone and a known number of employees of each type working in each zone. Implicitly there is a zone-to-zone matrix of home-to-work trips that can be estimated by standard techniques. Suppose now that the location of households in Table 1 and the location of industrial employment in the same table are taken as given. By first assuming that there will be one employee per household and then applying the Employee-Household Conversion Matrix given in Table 1, the number of industrial employees residing in each zone may be calculated. These were 146, 173, 98, 188, and 95, respectively. Note that the number of industrial em- ployeesworking in each zone is 150, 150, 400, 0, and 0, respectively.

It is possible to consider the proposal that each employee is to choose a place of work such that the total travel cost for all employees is minimized. If network link capacities, and the consequent congestion, are ignored for this illustration, this problem may be stated as a linear program problem. If there were specified link capacities (Y;j) for each link, which could not be exceeded, then another set of con- straints would be substituted for Equation 6. This new set of constraints would be of the form

$$V_{ij} \geq X_{ij} \geq 0 \quad (\forall\ i, j) \qquad (7)$$

Note that the objective function here is the same as that for the shortest path problem given in Equation 1. The constraints, Equations 5, are a set of flow-balance relationships similar to those of Equations 2.

Because the intrazonal travel costs are 1.0 for each zone, clearly the first consideration is that all employees residing in any zone first be assigned to jobs in that zone. Given the Archerville data, Zone 1 requires 4 workers and Zone 3 re- quires 302. Zones 2, 4, and 5 have 23, 188, and 95 surplus workers, respectively. The link flows produced by the simplex algorithm to solve this problem are shown in Figure 2. Note that if it were desired that some of the network links have a maximum allowable flow, then some constraints of the form of Equation 7 would have to be added.

Nonlinear Minimum-Cost Flow **Problem**

The minimum-cost flow problem described here can be recon- sidered as a nonlinear programming formulation. In the pre- viousformulation the linear objective function, Equation 4, was simply the sum of the trips (flows) on each network link times the travel cost of the link. The link costs were fixed, remaining constant regardless of link flows (though it was shownhowthelinkflowscouldthemselvesberestrictedbyuse of additional constraint equations). Suppose the more realistic view was taken that link costs depend on link flows. For the sake of illustration consider the following function, where link cost varies with linkflow

$$C_{ij} = C^{o}_{ij}\ (1.0 + \delta X^{2}_{ij}) \qquad (8)$$

where

$$\text{Min: } Z = \sum_i \sum_j c_{ij}\ X_{ij}$$

subject to

$Z\sum_j X_{ij}$ *ki*

k

O; if i = origin

——*D;* if *i* = destination 0 otherwise

(4)

(5)

(6)

$X_{ij} \geq 0 \quad (\forall\ i, j)$

C; = ”congested” or flow-related link travel cost,
= free-flow link travel cost,
X;j = link flow volume (trips), and
6 = a parameter.

With this function the link travel cost is equal to the free-flow cost when the link flow volume is zero. As link flow volume increases, link travel cost increases too.

In the linear version of this problem the solution involved only the finding of the minimum paths and the subsequent routing of trips along those paths. If there were specific link flow volume constraints, then the excess trips would be re- routed to the second shortest path. When link flows determine link costs the essential nature of the problem changes. The solution becomes a matter of adjusting volumes, observing the

whereO;isnettripsleavingnodeiand *D;*isnettripsarrivingat node *i*.

Equations 4, 5, and 6 are the general minimum-cost flow resulting costs, and then adjusting the volumes again. Thus, in a very real sense, even for the small problem size of the Archervilledata the complexity ofthe problem begins to defy

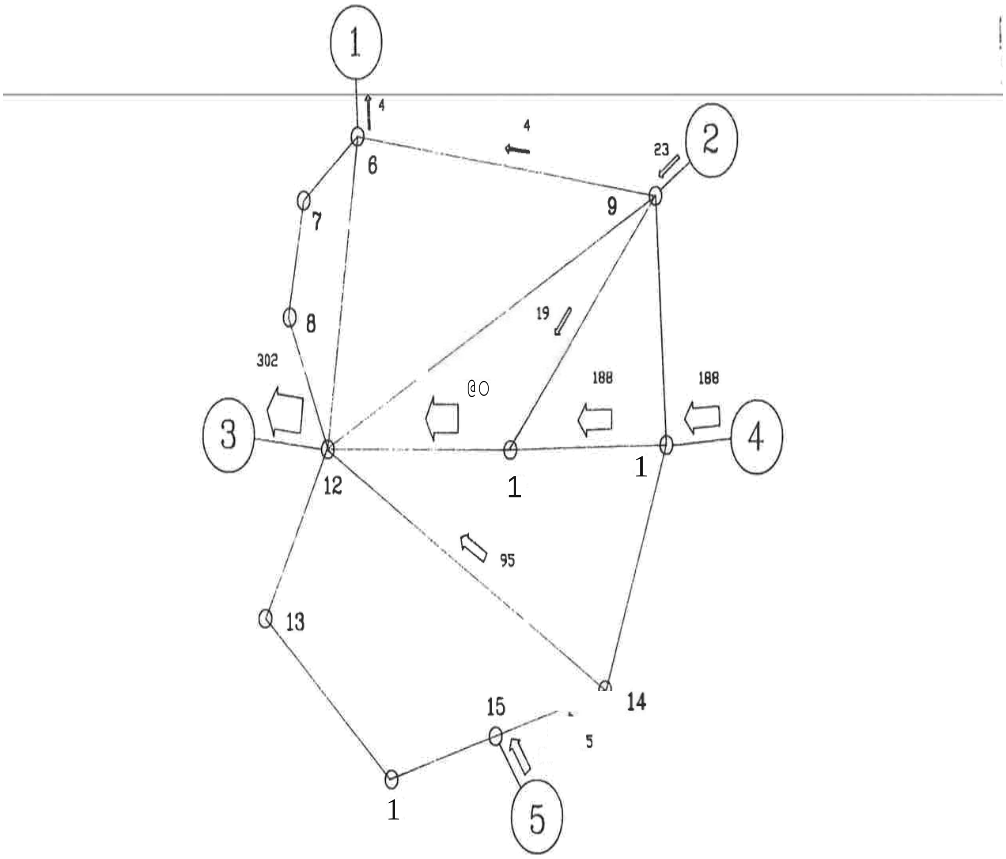

FIGURE 2 Archervllletrip flows resulting **from mlnimum-cost** Bow asslgnment**algorithm with constant llnltcosts.**

solution by inspection. The introduction of simultaneity or nonlinearity, or both, to a problem often transforms the prob- leminto one that lies beyond intuitive solution.

Incorporating Equation 8 into the original objective function of Equation 4 yields

k%m:Z=Z $\sum_{ij} \left[C^{o}_{ij} \ (1.0 + \delta X^{2}_{ij}) \ X_{ij} \right]$ (9)

and thusthe objective function becomes a cubic equation. The same linear constraints as before (Equations 5, the flow-bal- ancerelationships) still hold true, as do the nonnegativity constraints of Equation 6. This new set of equations is a nonlinear programming (NLP) problem with a nonlinear objec- tivefunction and linear constraints.

It was necessary toset a value for the parameter 6. Avalue

of 0.0002 was selected so that link flow volumes on the order

of 100 trips would result in a tripling of link cost. At this scale, link costs increase significantly, but not astronomically, with link flows on the order of those observed in the linear form of the problem described previously. The flows on the network that result from this new nonlinear problem are shown in Figure 3, and the link volumes, free-flow link costs, and con- gestedlinkcostsaregiveninTable2(onlythoselinks thathave flows are included). The results hold no great surprises butdo show a clear response to the reformulation of the problem so that link costs are a function of link flows.

It is interesting to compare the results shown in Figure 3 from the NLP solution with those in Figure 2 from the linear programming (LP) solution. The NLP solution, due to the effects of congestion on link travel costs, shows much greater utilization of network links. For the LP solution only 11 links were used whereas for the NLP solution 20 were used. As a result, thetripsonlinks X!0andXt0!2, whichwere 188in the LP solution, are only 112 in the NLP solution.

These examples only hint at the substantial additional work that has been done with user equilibrium and stochastic user equilibrium formulations of the traffic assignment problem as a mathematical program. Yet, they do give a clear way of seeing the assignment problem, as well as networks in general, ex- pressed in equation form. This will be particularly helpful in analyzing ways of linking transportation and location models. Thisinsight alsoprovidesamucheasierwayofcomprehending the problems of traffic assignment than did the “black-box” approach of traditional all-or-nothing assignment procedures. In the next section of this paper simple examples will be presented of location models presented as mathematical programs.

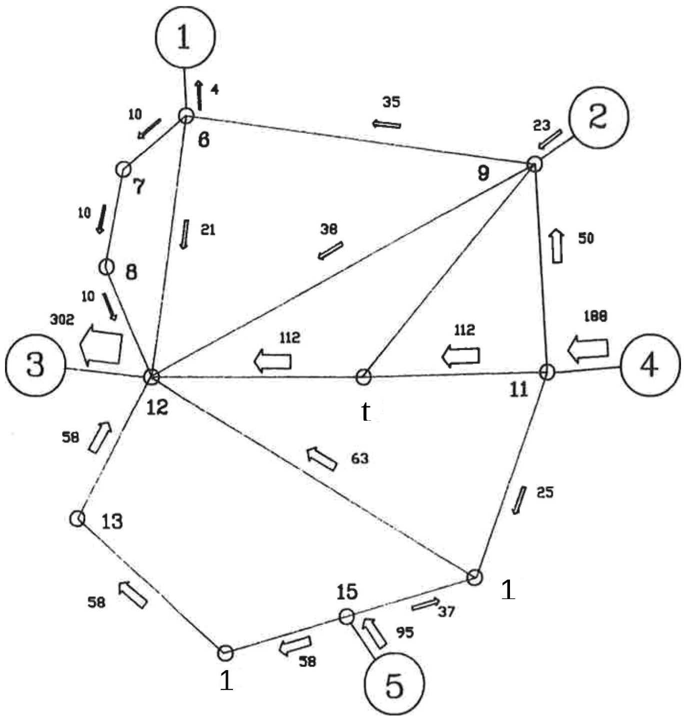

FIGURE 3 Archervilletrlpflows resulting from mlnlmum-cost flow assignment algorlthmwlt'avarlable link costs.

TECHNIQUE FOR THE OPTfifiAL PLACEMENT OF ACTIVITY IN ZONES: TOPAZ

TOPAZ is a mathematical programming technique that was originally proposed in the late 1960s and early 1970s *(9, 10).* The most complete discussion of the applications of the model is to be found in the book by Brotchie et a1. *(11).* The model was originally proposed as a method for determining least cost allocations of activities to zones. Perhaps the most recently publishedworkonTOPAZisbySharpe eta1.*(12)*fromwhich the formulation used here isadapted.

To begin, the model was abbreviated to a form for residence location only. Further, it was assumed, as is customary in these examples, that there is one employee per household and thus one work trip per household. The Archerville data show 0.55 low-income households per 1.0 employee and 0.45 high- income households per 1.0 employee. Thus the new, simplified, problem formulation becomes

Min: Z =Z ZZT_{ijl} c_{ijl} + $\sum_i \sum_j b_{ij}$ X_{ij}(0)

j l

subject to

$$\sum_l T_{ijl} - s_i X_{ij} = 0 \tag{11}$$

$$\sum_j X_{ij} =$$

$$\sum_i X_{ij} \leq$$

where

A_i

b_{ij}

c_{ijl}

s_i

r;

T_{ijl}

X_{ij}

Z_j

Z_j

A;

$T_{ijl} \geq 0, X_{ij} \geq 0$

= regional total of activityi,

= unit cost less benefit of locating activity *i*in zone j,

= unit cost less benefit of interaction for activity *i*between zone j and zone *1,*

= level of interaction (trips generated) per unit of activity *i,*

= trips attracted by employment per unit of activityi,

= level of interaction (trips) of activity *i* between zone y and zone *1,*

= amount of activity ito be allocated to zone j, and

= capacity of zone y.

(13)

(14)

(15,16)

Volume Cost Cost

$$\sum_{i} T_{ijl} - r_i X_l = 0 \quad (12)$$

Note that the second term in the objective function issimply a minimum-cost location term. The first term is thelinear

TABLE 2 ARCHERVILLE—COMPARISON OF FREE-FLOW AND CONGESTED LINK COSTS FOR NONLINEAR OBJECTIVE FUNCTION

Free-Flow Congested

$$\sum_{l} T_{ijl} - X_{ij} = 0$$

or, in words, the sum over all possible desiinationzones f ofall -trip»of-type i-leaving (4.e., produced in) xono y-nitistectualthe Link

$X_{2,9}$
$X_{5,15}$
$X_{-,-2}$ 21
$X_{7,8}$
$X_{9,6}$
$X_{10,12}$
$X_{11,9}$
$X_{11,14}$
$X_{14,12}$
$X_{15,14}$
$X_{15,16}$

	23	1.£D	1.11
*4.11	188	1.00	8.07
	95	1.£D	2.81
6,1	4	1.£D	1.(D
'6,5.00	10 5.44	2.€D	2.04
	10	2.ID	2.04
8,12	10	2.€D	2.04
	35	4.(D	4.96
xs,ia	38	7.00	9.07
	112	2.€D	7.06
	50	5.ID	7.52
11,10	112	1.ID	3.53
	25	4.€D	4.52
t2,3	302	1.€D	19.24
13,12	58	2.00	3.33
	63	6.€D	10.74
	37	2.€D	2.56
	58	2.€D	3.33
16,t3	58	4.£D	6.65

trips of type igenerated in zone y. In this example the trips of type *i*(i.e., household type i) generated in zone j equal the number of households of type i living in zone j.

Following the same reasoning, the r; will be equal to the household-type-per-employee ratios, so that the constraint equation(Equation12)willbeasgivenearlier,with*r;*--0.55, fi2=0.45,andX=totalemploymentinzonef.Then,inwords,

the sum over allpossible origin zones j of all trips of type i

terminating in zone *I* must equal the total households of type *i*attracted to employment in zone *1*. The total households of type *i*attracted to employment in I is simply the conversion rate times the employment.

An extensive seriesof test runs was done with thismodel

with varying weightings multiplied times one or another of the two components of the objective function. At the extremes, a location-cost-only solution and a transportation-cost-only solu-tionwere found. 'l'heminimum transport cost solution gave somewhat more dispersion of households to zones than did the minimum location cost solution. This was due in large part to the exogenouslydetermined location of employment. Were

“ transportation" problem. Taking the location problem first, it is clear that developing the necessary “ data" raises some difficult issues.Thenetbenefits*{b*j*;)*aresupposedtobethe net of the costs and benefits of locating a unit of activity type iin zone y. In many TOPAZ applications the virtual impossibility of Measuring bericfitsresulted in the b;, being simply a cost of location, to be iiiinirnized. For the Arclicrvillcexample several possibilities existed. The easiest way was simply to create an average annualized house cost variable, realizing that then the model would attempt to locate all households in the zone with the lowest house cost. All that prevents this location are the constraints on the amount of activity that can be accommodated in each zone.

Raisingtheissueof zonalconstraintsraises,inturn,theissue of converting activity types into land consumed. Here again, there were several possible ways to proceed: (a) regional land consumption rates by activity type, (b) zonal land consumption rates by activity type, or (c) exogenously developed housing stock estimates. For this illustration of the model a set of regional land consumption rates was assumed, and their values were set so as to allow all households to be accommodated by existing residential land in the region. The rates were 0.00525 land units per low-income (LI) household and 0.00646 land units per high-income (HI) household. The cost of location by zone was taken to be the average annualized house cost by zone, which was set to $7,800, $7,200, $7,600, $6,8€O,and

$9,668, respectively, for the five zones.

Next the transportation or interaction, cost term in the nhjer-- tivefunction (Equation 10) was examined This required the specification of data for the trip end constraints (Equations 11 and 12) as well. As mentioned previously, it was assumed that therewasonlyoneemployee perhouseholdandthattherewere

0.55 LI and 0.45 HI households per employee. Recalling that, in this example, only home-to-work trips are being dealt with and there is only one employment type, then hoths; will equal one, so that the constraint equations (Equation 11) will be

employment to be less dispersed, then, subject to thelandusecontraints, residential location would be less dispersedaswell.Another matter that will have interestingconsequencesforthe next set of tests, in which both locationandinteractioncosts are used along with a dispersion termto determineresi-dentiallocation, was that even though the locationcostcompo-nentwasalmostfourtimeslargerthanthetransportcostcom-ponent, the transport cosiportion of theobjectivefunction completely dominated the model solution. In severalmoretestruns the weighting of the location cost term in theobjective
function was varied from 0.01 to 2.00.

Over the range from 0.045 to 1.00 the location cost multi- plier resulted in a more or less gradual shift from the transport- cost-only solution toward the location-cost-only solution. From a value of slightly less than 1.0 to a value of 1.1852 the multiplier causes no change in the model solution. At a value of 1.1855 the multiplier results in a model solution identical to that of the location-cost-only solution. Thusat some critical point whcrcthcmultiplier of the location cost term in the objective function is between 1.1852 and 1.1855, there is a sudden shift in the model solution from an apparently stable intermediate solution to the location-cost-only solution. Al- though it is presented as an aside here, clearly the matter of model sensitivity and solution stability is an area for future research. In any case, for a midrange weighting of 0.70, the results of the TOPAZ model solution were as given in Table3.

isumericainxampieincorporating Dispersion

The location patterns produced by the linear programming version of this model, particularly when examined at the zone- to-zone trip level, are rather lumpy. The addition of a nonlinear dispersion term to the objective function can make a noticeable difference. This is achieved by substituting a constrained grav- itymodel for the linear “transportation" model portion of the objective function.

TABLE3ARCHERVIL OVATION OF HOUSEHOLDS BASED ON LOCATION COST PLUS

TRANSPORTCOSTCOMPONENTSOFMODEL(=0.70)

Households				Households, Showing Place ofWork	
Zone	LI	HI	Total	LI	HI
1		0294	294	0	1-135 3-69 4-45 5-45
2	359	173	532	1-166 2-193	2-157 4-16
3		0155	155	0	3-155
4	441		28469	3-276 4-110 5-55	4-28
5		0	0 0	0	0

Now: = locati‹xicost multiplier in objectivefunction.

The standard doubly (or fully) constrainedspatialinteraction model has the form

$$\text{Min: } S = \frac{1}{\beta} \sum_i \sum_j \sum_l T_{ijl} \ln T_{ijl}$$

$$\sum_j$$

T j ---- *Aj B 0 j D* exp(--|3cj) where

Tj -- number of trips between zone j and zone *1,*

Oj= trips generated in(originating from) zonej,

D = trips attracted to (terminating in) zonef,
c; j -- zone-to-zone travel cost, and
§ = aparameter;

(17)

$j \quad l$

subject to

$\sum_{l} T_{ijl} - s_i X_{ij} = 0$

$_{j} \quad T_{ijl} - r_i X_l = 0$

s
+ZZZ
$T_{ijl} \; c_{ijl} + \sum_{i} b$ *ij ij*
(24)
(x)
(26)
and where
Aj---- [Z *B D* exp(Qcjt)]°'

B -- [$\sum_{j} A_j \; O_j \exp(-\beta c_{jl}) \Big]^{-1}$

(18)
(19)
$\sum_{j}$=*A;*
s
$T_{ijl} \geq 0, \; X_{ij} \geq 0$

X_{ij}

$X_{ij} \leq Z_j$

27)

(28)

(29,30)

IthasbeenshownbyMurchland*(13j*andWilson*(14j*thatthis model can be derived from an equivalent optimization prob- lem. The form of that problemis

The Archerville data were again used for tests of the model, which now required a numerical value of Q. This, however, raises an interesting question, which relates directly to the previously described experiments with weightings or multi- pliers of the terms in the objective function. To simplify the

Max: *S* = -- $Z_{j} \sum_{l} T_{jl} \ln T_{jl}$
subject to

(20)
coming discussion it will be convenient to think of the objec- tive function, Equation 24, as having three components:

$$\sum_l T_{jl} = O_j \quad (21)$$

$$\sum_j T_{jl} = D_l \quad (22)$$

$$\sum_j \sum_l c_{jl} T_{jl} = C \quad (23)$$

where the only new value is C, taken to represent the total system travel cost.

There is a relationship between § in Equations 17--19 and C in Equation 23. The |3 in the spatial interaction model produces a dispersion of trips away from the optimum or minimum-cost solution. What human behavior might account for this disper- sionis unspecified but presumably includes such factors as variables not in the model, as well as variations in individual perceptions of **costs** and differences in individual utility functions.

To introduce the spatial interaction model into TOPAZ in lieu of the minimum-cost " transportation" model requires that the model of Equations 20--23 be substituted for the transport cost term in the TOPAZ objective function. Following al- gebraic manipulations, this yields (after changing signs to al- low minimization)

$$\text{Min: } Z = \lambda_3 U_3 + \lambda_1 U_1 + \lambda_2 U_2 \quad (31)$$

where k;,, and are arbitrary weights, and where U is the transport cost term, U is the location cost term, and U is the entropy term. The no-dispersion solution given in Table 3 was for k equals one, equals 0.70, and equals zero.

Thevalueof as discussed here is the inverse of the § of Equation 17 and thus will directly affect the extent to which location is dispersed from the no-dispersion casewhere = 0 and thus |3 = -. With a of 20 and thus }3 = 0.05, the dispersion shows quite clearly in the results given in Table 4. A further increasein , to 30, giving Q = **0.033,** yields even greater dispersion. Numerous tests were run with different combinations of values fork, ,and . In all cases increas- ingvaluesof produced increased dispersion of household location. The most dispersion was achieved with values of k equaltoone, in thevicinity of0.3,and equal to 30 or more

TABLE4 ARCHERVIL LOCATION OF HOUSEHOLDS BASED ON LOCATION COST PLUSTRANSPORTCOSTCOMPONENTSOFMODEL(=0.70), INCORPORATING DISPERSION, WITH g = 0.05

	Households			Households, Showing Place ofWork	
Zone	LI	HI	Total	LI	HI
1		45258	302	1-42 3-3	1-128 2-21 3-84 4-14 5-11
2	319	205	524	1-105 2-181 3-25 4-4 5-4	1-6 2-135 3-17 4-32 5-15
3		20138	158	1-1 3-19	3-117 4-12 5-9
4	416		49465	1-1g 2-11 3-230	3-6 4-33 5-10 4-106 5-51
5		0	0 0	0	0

None: = location cost multiplier in objectivefuncuon.

COMBINED **ACTIVITYLOCATION**ANDTRIP $\sum_l T_{ijl} - s_i X_{ij} = 0$ASSIGNMENTMODEL

$$\sum_j T_{ijl} - r_i E_l = 0$$

s

Having shown in the preceding two sections of this paper how both a traffic assignment model and an activitylocationmodel $\sum_{j} X_{ij} = A_i$ could be formulated as mathematical programming problems, it isnow possible to consider linking them togetherintoa $\sum_{i} X_{ij} \leq Z_j$ single model to solve both problemssimultaneously.

First, note that the original transport cost term in th $T_{ijl} \geq 0,\ X_{ij} \geq 0,\ d_{mn} > 0$ tive function is simply the sum of all trips between

(34)

(35)

(36)

(37)

(38, 39,40)

origins j and I, times the cost of those trips. It is reasonable to assume that, regardless of whether a congested or uncongested networkisbeingusedinthemodel, the**costs**usedwillbethose of the minimum paths through the network. In the location model the zone-to-zone costs used are the exogenously deter- mined minimum **costs** given as input to the calculations. The minimum-cost flow problem determines the set of minimum cost paths through the network. Given that these minimum paths are the same, the flow cost produced by the minimum- cost flow problem solution should be identical to the transport cost term in the location modelobjective function.

In the nonlinear version of the location model the entropy term of the cost function takes care of keeping the zone to zonetrips in the objective function, thus the transport cost term can be replaced by the objective function from the minimum-cost flow problem. It will, of course, be necessary to add the flow balance constraints in order to describe the network, and to set them equal to what is now a variable set of trip origins and destinations. These are now variable because the activity loca- tions, and thus the trip matrix, are to be determined as part of the model solution. Thusthe combined model has the follow- ing form:

$$\text{Min: } S = \sum \sum \sum d_{mn} F_{imn}$$

$$T_{ijl} \ln T_{ijl} + \frac{1}{\beta} \sum_{i} \sum_{j} \sum_{l}$$

subject to

$$\sum_{i} \sum_{n} F_{imn} - \sum_{i} \sum_{n} F_{inm} = \begin{cases} \sum_{i}(s_i X_{im} - r_i E_m) & m \leq J \\ 0 & m > J \end{cases} \quad (33)$$

where the variable names d and *F* (link cost and link flow) are substituted for the c and X used in the minimum-cost flow problem equations. I isthe number of zones (load nodes). In additionemployment in zone *I* is represented by £ in Equation 35 to avoid confusion.

The objective function now has three terms, minimum-cost flows, eotropyor trip dispersion, and minimum-cost location. Anewsetofconstraints(Equations 33)isaddedtotheprevious set of location model constraints to incorporate the flow-bal- anceportion of the minimum-cost flow problem. This being done it is possible to solve the combined problem for both link flows and household location. Now it is possible to include congestion in this rnblemtrim. The first term in the ohjectivefunction (Equation 32) is modified as per Equation 9 to make link cost a function of link volume (flow). The remainder ofthe objective function and all of the constraints areunchanged.

CONCLUSIONS

The work described in this paper is only a brief introduction to the topic, yet certain conclusions can now be clearly drawn. The first of these is that numerous model tests using the kinds of models described here indicate that linear mathematical programming models of location are inherently unrealistic. The least-cost zone will get all possible locators even it the next-to- least-cost zone is only marginally more expensive. The objec- tivefunction component weighting problem implies that an arbitrary difference in units of measurement (e.g., between hundreds of dollars or thousands of dollars for annualized rent) can result in one component of a model solution's being domi- imnvoverrunner.

To a rather considerable degree theconstraint equations of a linear programming model can ameliorate some of thesedif- faculties. This is, however, a mixed blessing. That theycan ameliorate some of the difficulties also points up the rather considerable extent to which they determine the model solution. The constraints (e.g., available residential land per zone) must be exogenously determined, yet their determina- tion, in and of itself, implies a difficult forecasting problem. The availability of data for constraints during the development of a mathematical programming location model can give afalse sense of confidence in the model's predictive power if the problems of forecasting the constraints are not taken into account.

Another major **issue is** that of obtainingthe necessarydata. Housing costs are notoriously difficult to estimate.WiththeTOPAZ model the literature suggests that it has oftenbeensodifficult to estimate the benefits of location orinteractionthatonly costs were used. Yet, both versions ofTOPAZwouldcertainly have given different results if somelocationaladvan-tagevariable had been used to yield a "net"locationcostvariable for the location cost temiin theobjectivefunction.Finallythere is the computational problem.TheGAMSpackage *(15}* was used for all of the Archervilletestsreportedhere. The linear fomiulationswere run on an IBMPCwith640K of memory and a hard disk drive. These runs tmkjustafew minutes each. The nonlinear formulationswereproblem-aticon the PC; some took 2 or more hours to solve,andotherswouldnotrunatall.Finally,aversionofGAMSfortheIBM3081 GX mainframe was used for thenonlinearformulations.In the combined model there were 622 variables intheobjec- tive function. The objective function for thefinalnonlinearversion of TOPAZ, alone, had 110 variables, but, iftherewerea 30-zone region to be analyzed, the objective functioninthenonlinear TOPAZ model would have 3,660 variables. Thisisarather sizable problem, and yet a 30-zonespatialinteractionmodel is really too highly a8gregated for most**policy**analysispurposes. Further, the examples presented usedonlytwohousehold types and no housing stockconsideration.Currenttransportation and location modeling applications tendtohavefrom 200 to 300 zones. With no increase in activitytypes,250zones would yield 250,5tD variables in TOPAZ.Themodelformulations that combine location and tripassignmentareeven larger. Problems of this size are reallyquiteimpracticalfor direct solution. The problems of solvingoptimizingmodelsof a realistic (from the applications point of view)size canbe

dealt with by decomposition procedures if it is desirabletomaintain the mathematical programmingformulations.Thepossibility of transforming programming modelsintospatialinteraction models as briefly mentioned earlieroffersanotheravenue of approach, Both of these approaches will beexam-
inedin future research.

Despite these cones acre are several important points to belearnedfromtheseexperiments.Perhapsthemostimportant is that developing these model formulations and then testing theirbehaviorgiveswonderfulinsightintovarioushypotheses
such formulations became clearly evident in these experiments. At the same time the experiments clearly illustrated the need for inclusion of dispersion terms in such models. The inference is that although locational behavior may be said to be, in principle, an optimizing process, in actuality there are ob- viouslyother factors that result in a dispersion of locations around a simple “least-cost” optimYet the optimizing process provides a model-building rationale that can be par- ticularlyhelpful in understanding the implications of model structure and can thus, in turn, be expected to improve model- ingpractice aswell.

REFERENCES

1. S. Putman. FutumDirections for Urban Systems Models: Son Pointers from Empirical Investigahons. In Advances in (/rban*Systems* ñf*odnlling*(B. Hutchison and ML Batty, eds.), North-Holland. Amsterdam. The Netherlands,1986.
2. S.Putman.*IntegratedUrban*3f*odelz.*PionLtd.,London,England,1983.
3. M. Ben-Akiva and S. Lerman. *Discrete Choice Analysis.* MIT Press, Cambridge, Mass.,1985.
4. Y.Sheffi.*UrbanTransportationNetworks.*Prentice-Hall,Inc., Englewood Cliffs, N.J.,1985.

S. A. Anas. *Residential* Lnc‹zfiozt*Markets and Urban Trancportation.*
Academic Press, New York,1982.

6. A.Wilson,*I.*Coelho,*S.*Macgill,andH.Williams.*Optimizationin LocationalandTransportAnal:ysis.*JohnWiley&Sons,Chiches-ter, England,1981.
7. D.**Boyce**andR.Eash.Applicationofa**Combined**Residential Location,Mode,andRomeChoiceModeitoSDategicPlanningfor the Northeastern Illinois Regional Transportation Authority.Presentedat66thAnnualMeetingoftheTransportationResearch Board, Washington, D.C.,1987.
8. P. **Prastacos.** Urban **Development Models for the** San**Francisco Region:**FromPLUMtoPOLIS.Presentedat64thAnnualMeet- ingof the Transportation Research Board. Washington, D.C., 1985.
9. J. Brotchie. A General Planning Model. 3fnnagemenf*Science,* Vol. 16, 1969, pp. 265-266.
10. R.SharpeandJ.Brotchie.AnUrbanSystemsStudy.JtynfAus- *tralian Planning Institute Journal,*of. 10, 1972, pp.105—118.t1.I.Brotchie,J.Dickey,andR.Sharpe.*TOP*AZLec*tazn*NotesinEconomics and Mathematical Systems 180.Springer-Verlag,
Heidelberg, Federal Republic of Germany, 1980.
12. R.Sharpe,B.Wilson,andR.Pailot.ComputerUserManualfor ProgramTOPAZ82.DivisionofBuildingResearch,Common-wealth Scientific and Industrial, Research Organisation, Highett,Victoria, Australia,1984.

13. *I.*Murchland.*SomeRemarksontheGravityfodelofTripDis-IribizfionandanEquivalentMaximizingProcedure*.LSE-TNT-38. London School of Economics, London, England,1966.
14. A. Wilson. A Sta0stical Theory of Spatial DistribufionModels. *Transport‹zfion Researc:h.*Vol. 1, 1967, pp. 252-269.
15. D.KendrickandA.Merraus.*GAMS,An*iztfroduciion.TheWorld Bank, Washington, D.C.,1985.

aboutlocationalbehavior.Theeminentsensibilityofdescrib-

inglocational behavior as an optimizing process is beyond reproach Theeffects, and general importance, of constraints in