Michael F. Goodchild, Brian Klinkenberg
and Donald G. Janelle

A Factorial Model of Aggregate Spatio-Temporal Behavior: Application to the Diurnal Cycle

The cross-sectional nature of much social data, coupled with the static view provided by maps and current spatial data handling software, have produced a tradition of research on urban spatial structure that is largely two-dimensional and derived from residential locations. The paper presents an analysis of a space-time diary data set collected in Halifax, Nova Scotia. A series of transformations are used to convert the individual diary records to a three-mode matrix of intensities, which is then analyzed using the PARAFAC three-mode factor model. Home / work is found to be the strongest organizing dimension of the urban space-time, followed by entertainment, shopping, and education / work. We show how these dimensions appear to varying degrees in different locations, time periods, and human activities. The paper argues for a dynamic view of urban spatial structure in which only the physical facilities remain static.

INTRODUCTION

Maps are static, two-dimensional representations of geographic phenomena, and therefore of necessity better at capturing the static dimensions of geographic distributions. Although digital geographic databases are not limited by the same technological constraints (Goodchild 1988; Marble 1990), they remain primarily stores of digital representations of maps, and thus similarly static and two-dimensional (Langran 1989, 1992; Langran and Chrisman 1988; Raper 1989). Recently, Couclelis (1991) has commented that current GISs are also limited in their views of space (Gatrell 1983; Sack 1980). The map reflects a container view of space, and records faithfully the absolute locations of objects. However, almost all of our understanding of social process is built upon a relative view of space, in which

The National Center of Geographic Information and Analysis is supported by the National Science Foundation, grant SES 88-10917. The authors also acknowledge the financial support of the Social Sciences and Humanities Research Council of Canada and the University of Western Ontario. They are grateful to Richard Harshman who provided access to PARAFAC, and to the organizers of the Dimensions of Metropolitan Activity (DOMA) project, particularly Andrew Harvey, for access to the Halifax data. They also thank Harry Taylor, Jr., William Sauruk, Micheal Dal Bello, and Terry van Kessel for research assistance.

Michael F. Goodchild is professor of geography at the National Center for Geographic Information and Analysis, University of California, Santa Barbara. Brian Klinkenberg is assistant professor of geography at the University of British Columbia. Donald G. Janelle is professor of geography at the University of Western Ontario.

Geographical Analysis, Vol. 25, No. 4 (October 1993)

interaction plays a much greater role than absolute location [there are echoes of this point in the familiar site/situation dichotomy and in Haggett's distinction between location and place (Haggett 1979)]. In summary, it is contended that analysis of social process requires a time-dependent, multidimensional, relative space rather than the static, two-dimensional, absolute view embedded in current GIS technology.

The purpose of this paper is to examine this contention in the setting provided by a major study of the time geography of a Canadian city (Janelle and Goodchild 1983a, b; Goodchild and Janelle 1984; Janelle and Goodchild 1987; Janelle, Goodchild, and Klinkenberg 1988). The term "time geography" was in use in Sweden in the mid-1960s, but was first used in English by Hägerstrand (1970) in a discussion of the dynamic element in human spatial behavior (see also Parkes and Thrift 1980). For example, time dependence on a scale of months or years underlies studies of migration; on a scale of minutes and hours it underlies studies of commuting, shopping, and other aspects of daily human space-time behavior [for an interesting discussion see Holly (1978)]. Previous papers have described earlier analyses, particularly those concerned with the diurnal geography of the city. In this paper we discuss the modeling of space-time behavior both at the individual level, as trajectories in three-dimensional space-time in the style of Hägerstrand, and also at the aggregate level, as time-slice replications of urban ecology. We discuss some of the issues presented by the need to handle these types of data, particularly given the previously described limitations on current geographical data handling software.

While the context of the paper is provided by the literature on empirical studies of space-time behavior and factorial ecology, and by the issues of spatial data representation and handling embedded in GIS, it is also appropriate to develop relationships between this work and the literature on urban spatial dynamics (Pumain 1991). In recent years there has been substantial interest in housing dynamics and their impact on urban structure (Clark, Deurloo, and Dieleman 1984; van Wissen and Rima 1988; van Wissen, Rima, and Nijkamp 1986), particularly on linkages between household demographic processes and spatial behavior. Another class of theoretical models of the urban system is based on extensions to the land use theories of von Thünen (see, for example, Fujita 1989) and deals with competition for space between residential and industrial land uses, coupled with factors of transportation and congestion. Dynamic models of shopping behavior have been developed by Wrigley and others (Hauer, Timmermans, and Wrigley 1989; Wrigley 1988) to describe spatial behaviour within a fixed and known urban structure. Finally, the relative merits of longitudinal and cross-sectional analysis for understanding dynamic spatial behavior have been discussed by Pickles and Davies (1989).

The first section of the paper describes the Halifax time geography project and its diary and Census data sets, and outlines the overall objectives of the project. The second section discusses some of the issues encountered in managing and manipulating the data. This is followed by a review of available models for aggregate space-time data, and by a discussion of the application of one particular model, PARAFAC (Harshman 1976; Harshman, Ladefoged and Goldstein 1977).

THE HALIFAX TIME GEOGRAPHY PROJECT

Despite its obvious and widely publicized limitations, the Census remains the major source of information on the geographical distribution of population in most countries. Our empirically based understanding of social processes is therefore limited by two important problems. First, although most activities and expenditures occur outside the home, the location of individuals is determined solely by place of residence. A second and related issue is that much geographical research is thus restricted to the night-time locations of the population, and unable to focus on day-time locations, or diurnal movement.

Between October 1971 and March 1972, a study was conducted of the activities and movements of 2,141 people in the Halifax Census Metropolitan Area, Nova Scotia. Sampling occurred in a randomly chosen subset of sixty-eight of the 380 Census Enumeration Areas of the Halifax CMA. Respondents recorded what they were doing, where, and with whom for twenty-four hours. The data set used in the analysis reported here covers only the 1,561 respondents who were living in the cities of Halifax and Dartmouth; who recorded activities on the weekdays, Monday through Friday; who were aged from nineteen through sixty-four; and where at least one person was employed in the household. Activities were expressed in terms of ninety-nine categories, and geocoded to 100-meter accuracy. All respondents were matched with full Census returns from the 1971 Census of Canada [see Elliott, Harvey, and Procos (1976) for an overview of the data collection, and Cosper and Shaw (1984) for an assessment of the data set].

The general objective of the Halifax project has been to describe and model respondents' movements and activities, and to generalize these to statements about the population as a whole. In the Hägerstrand view, individual space-time behavior can be understood in terms of constraints—economic, social, biological, technological, or geographical. At the aggregate level, the data allow us to build up a picture of the city not only at night, but as it changes through the diurnal cycle, as people move from residence to place of work and to locations of shopping and entertainment. Instead of the static, residential view, we see the city as a container within which its population moves continuously. Patterns build and decay throughout the daily cycle, as people congregate and disperse. Socioeconomic variables such as wealth are continuous attributes of individuals but only transient attributes of places; as people move, they carry many of their socioeconomic characteristics with them. Much human spatial behavior is determined at least in part by socioeconomic aggregations, such as occur in many residential neighborhoods. Thus we envision urban ecology as a dynamic system, created by the movement of individuals and at the same time responsible for many aspects of that movement. We are interested in knowing how urban ecology varies during the day, and whether it is stronger at certain parts of the diurnal cycle than others.

DATA MANAGEMENT ISSUES

In principle, the Halifax data set identifies the location and activity of any respondent at any time during the sample period. The original data contain approximately 65,000 such records, giving the start time and activity class for each activity event reported. Thus the data set forms a collection of approximately 65,000 points in three-dimensional space-time, with a spatial resolution of 100 meters and a temporal resolution of 1 minute. By sorting, we can link the records of any one individual into a sequence of activities, and thus into a line or trajectory in space-time. Vertical lines indicate static activities; for activities movement, such as travel to work, the corresponding lines are oblique. However, since the data contain no information on the route of movement, a straight line in space-time provides only a crude approximation to actual location, as it assumes the ability to travel in straight lines at constant speed. Nevertheless it would be possible to compute a point distribution of the respondents at any given time during the day, as a cross-section of space-time (Hägerstrand 1970).

TABLE 1
Average Durations and Representative Hours of Aggregated Classes of Activity

Activity	Duration (minutes)	Hour of highest involvement
Sleeping	416.5	1:00 A.M.–2:00 A.M.
Traveling to work	15.7	8:00 A.M.–9:00 A.M.
Working before noon	76.7	9:00 A.M.–10:00 A.M.
Lunch	31.0	12:00 P.M.–1:00 P.M.
Working after noon	64.9	2:00 P.M.–3:00 P.M.
Traveling after work	15.7	5:00 P.M.–6:00 P.M.
Early evening discretionary	57.2	6:00 P.M.–7:00 P.M.
Late evening transfer to bed		10:00 P.M.–11:00 P.M.

Although it provides a precise description of locations and activities, the representation of the data set as a collection of 1,561 trajectories in space-time, with over 65,000 straight line segments, is far from a convenient basis for analysis or modeling. Instead, we aggregated the data into measures of intensity in the three modes of time, location, and activity.

To simplify the temporal dimension, we undertook an analysis of the frequency of activities occurring at any time of day. Table 1 shows the eight hours most representative of various aggregated classes of activity. For example, the period 1:00 A.M.–2:00 A.M. is when the greatest proportion of the sample reported sleeping. The mean duration of each activity is also indicated—respondents indicated 416.5 minutes sleeping on average. 10:00 P.M.–11:00 P.M. is the hour when the largest number of respondents reported transfer to bed, although this was not identified as a specific activity in the survey. Note that the mean duration for working activity before noon, for example, is defined as the total time reported for that activity event by all respondents for whom the activity event occurred at least in part in that hour, divided by the total number of respondents: we would expect this figure to be substantially less than the average number of minutes worked before lunch. For each of the eight periods shown in Table 1, we tabulated the numbers of respondents involved in each type of activity. Some respondents were counted more than once, if they reported more than one activity in the period.

Aggregation of the spatial dimension is more difficult because of the stratified sampling design, which selected respondents having residential addresses in only sixty-eight of the 380 Census Enumeration Areas. Although the respondents visited a much larger sample of EAs during the day, at no time did their geographical distribution approach uniformity. For analysis, we aggregated the data into geographical reporting zones satisfying the following criteria:

1. Each zone contained at least one of the sixty-eight sampled EAs;
2. The number of zones was approximately equal to the number of Census Tracts in the Halifax CMA;
3. The zones were constructed from contiguous aggregations of EAs;
4. The zones contained equal shares of the sample of respondents in each of the eight time periods defined in Table 1;
5. The zones were compact and singly bounded; and
6. Residential populations were homogeneous across zones on the seventeen selected demographic variables shown in Table 2.

These contrived reporting zones were termed pseudo-Census Tracts or PCTs. All six requirements cannot be satisfied simultaneously, so the twenty-nine zones shown in Figure 1 represent a compromise, particularly between requirements 4 and 6 above.

TABLE 2
Demographic Variables Used to Ensure Homogeneity in Building Pseudo-Census Tracts (PCTs)

AVGCH	Children per household
AVGPPH	Persons per household
AVGRPD	Rooms per household
AVGRENT	Rent per household
AVGINC	Income per household
PFEM	Percent female
PWDS	Percent widowed, separated, or divorced
PSIN	Percent single
FAMHUNV	Percent with university-educated family head
ANGRD	Percent Anglican
RCJRD	Percent Roman Catholic or Jewish
DWELOW	Percent of dwellings owned
WFEMP	Percent employed
DWELSD	Percent in single detached homes
POP2059	Percent 20 to 59 years old
DWELNOAT	Percent households without autos
WFMANCON	Percent blue collar labor

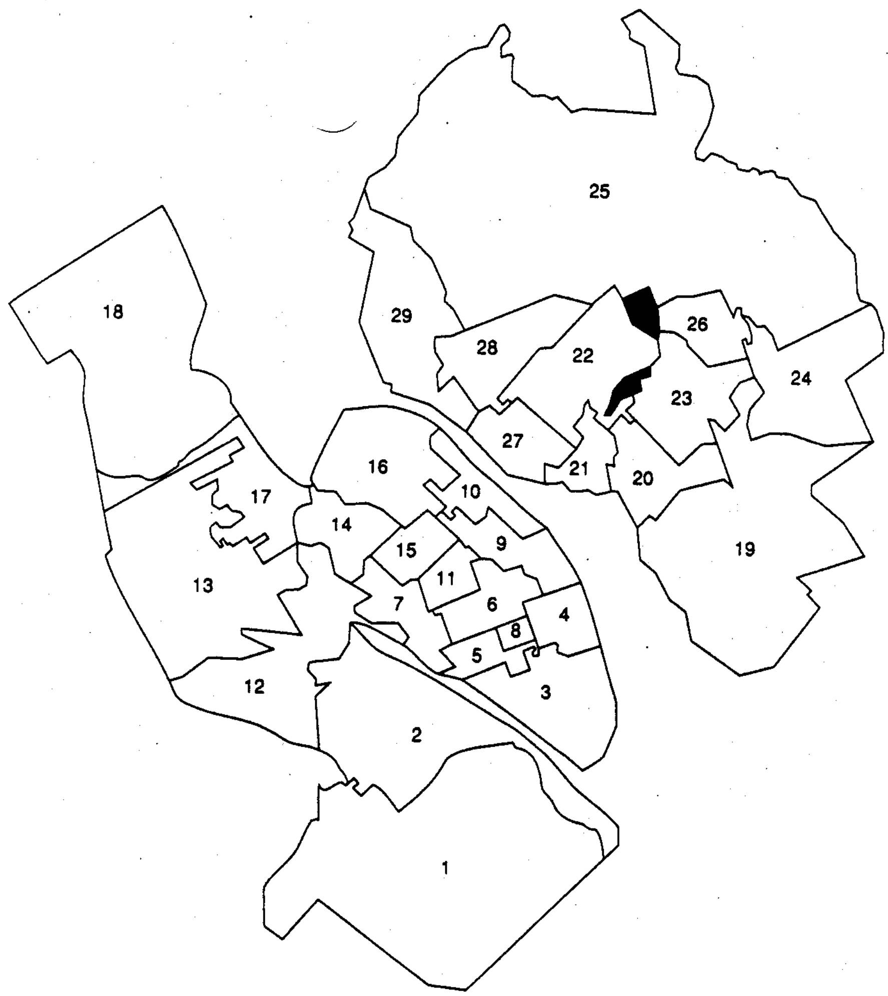

FIG. 1. Outline Map of the Pseudo-Census Tracts (PCTs) Used for Analysis. Each zone is approximately homogeneous on fourteen census variables, compact and singly bounded, and containing approximately equal numbers of respondents in each time period

TABLE 3
Major Features of the Halifax/Dartmouth Metropolitan Area and Corresponding PCTs

Halifax CBD	4, 9
Dartmouth CBD	27
Major shopping centers	12, 14, 23
Universities	5, 6
Hospitals	8
Low income residential	10, 16, 22
Military installations	16, 19, 20
Port installations	3, 17

To construct the PCTs, twenty-nine EAs were first randomly selected from the sixty-eight sampled EAs to act as zone cores [see Rossiter and Johnston (1981) and Goodchild and Hosage (1983) for similar procedures.]. Adjacent EAs were then added one at a time, so as to produce the minimum increment in the total within-group sum of squares on the seventeen demographic variables, until all EAs had been added to a zone. Because this procedure gave no weight to criterion 4 above, a process of manual adjustment, reassigning EAs from one zone to another without violating the contiguity constraint 3 above, was used to create an improved balance of populations in each zone and time period.

The city of Halifax is centered on a slope immediately adjacent to an arm of the Atlantic, and has spread to include extensive suburbs. The city functions as the political, administrative, and commercial capital of the province, and also includes major concentrations of medical and postsecondary education facilities. Across the arm lies the city of Dartmouth, which includes many of the military installations as well as suburbs, and a small CBD. Major geographic features of the urban area and their corresponding PCTs are shown in Table 3.

For each PCT in each of the one-hour time periods shown in Table 1, we computed the characteristics of the population present by identifying those respondents who reported being located in the PCT during that period. From the activity records, we calculated the numbers of respondents involved in various activities, weighted by the proportion of the hour spent in each activity. In addition, we identified the numbers in the various "with whom" and "site" categories shown in Table 4. In total, nineteen activity categories, ten "with whom" and eight "site" categories were tabulated for each of the twenty-nine PCTs and eight time periods. The result is a three-dimensional ("three-mode") matrix of intensity scores for eight times, twenty-nine location zones, and thirty-seven activity-related variables (termed simply "variables"). This $8 \times 29 \times 37$ matrix was the basis for the analysis and modeling reported in this paper. In what follows, we use the subscript i ($i = 1, \ldots, 29$) to refer to locations, j ($j = 1, \ldots, 37$) to variables, and k ($k = 1, \ldots, 8$) to time periods. x_{ijk} denotes an element of the data matrix, an estimate of the probability that a randomly chosen respondent located in PCT i in time period k would have reported being involved in activity j at a random time in that one-hour interval, found by dividing the total time reportedly spent by respondents in activity j in PCT i during time period k by the total time spent by all respondents in PCT i during time period k. Note that a single respondent can contribute to the totals for more than one activity in any one time period, and to the totals for more than one PCT.

THE MODEL

The proposed model adds the temporal dimension to traditional perspectives on urban ecology. We conjecture that diurnal variation in the distribution of population over the city, and their activities, can be described through a small number of

TABLE 4
The Thirty-Seven Variables Used in PARAFAC Analysis

ACTCOD1	Work income-producing time
2	Work nonincome-producing time
3	Travel to and from work
4	Household work
5	Household maintenance
6	Child care
7	Marketing and shopping for household
8	Other household work
9	Travel related to household
10	Sleep and personal care
11	Meal times
12	Educational time
13	Recreation outdoors and sports
14	Entertainment away from household
15	Leisure: socializing
16	Leisure: radio and TV
17	Reading, noneducational
18	Other leisure time
19	Travel for leisure only
WHOM0	Alone
1	With spouse or fiancé
2	With child of household
3	With adult of household
4	With relative
5	With colleagues
6	With organization and club associates
7	With neighbors
8	With nonbusiness associates
9	With others and unknown
SITE0	At home
1	At work place
3	Someone else's building
4	Outdoors travel by public transit
5	At public or private building
6	Leisure indoors
8	Restaurants and bars
9	All other places

basic, underlying dimensions. These dimensions are inherently unobservable, but appear in each location (PCT), time period, and activity to a degree that is unique to that location/time/activity combination. The effects of location, time, and activity may be to some extent separable; underlying dimensions may appear more strongly in some time periods, or in some locations, or in some activities. The objective of the analysis will be to determine the nature of each underlying dimension, and the degree to which it appears in each location, time, and activity.

A variety of approaches to this problem have been discussed in the literature. Taylor and Parkes (1975) restructured their three-mode data to two modes by regarding time periods as additional locations. In our case, the $8 \times 29 \times 37$ matrix would become 232×37. Their analysis was conducted on artifically constructed data using factor analysis, with the following model:

$$x_{lj} = \sum_{p} a_{jp} F_{lp} + U_{lj} \tag{1}$$

where l ($l = 1, \ldots, 232$ in our case) denotes the location-time combinations in the renumbered rows of the matrix. In this factor analytic model the Fs are the (unknown) underlying dimensions ("factors") of variability, each F_{lp} denotes the amount of underlying dimension p ($p = 1, \ldots, s$) present in each location-time combination, and the a_{jp} or "loadings" denote the amount contributed by factor p

to variable j. The Us or "uniqueness" measures represent the degree to which each variable j is unique, rather than representative of underlying dimensions. This model represents an interesting extension of conventional factor analysis to the time-dependent case [for a review of the use of factorial methods in the analysis of static, residence-based data see Davies (1984]. But despite its simplicity, because time periods appear as additional locations, it follows that information on the behavior of one PCT over time periods, or of geographical variation in any one period, is lost in this approach.

An alternative approach would be to analyze each time "slab" separately, by using eight replications of two-dimensional analysis. The model in this case would be

$$x_{ijk} = \sum_p a_{jpk} F_{ipk} + U_{ijk} \tag{2}$$

yielding independent sets of factors in each time period k. Because of the inherent problems of comparing certain types of factorial solutions (see, for example, Berry 1971), it would be necessary to adopt one time period as a standard, and to rotate all other solutions to it using one of the so-called Procrustean methods before comparison could take place. The results of Procrustean rotations are known to be difficult to interpret (Berry 1971). Moreover, results would be invariant under rearrangements of PCTs within time slices, and thus insensitive to the behavior of specific PCTs through time.

Both of these methods reduce the three-mode matrix to two modes, and in doing so prevent the identification of factors operating in all three modes. By contrast, methods of three-way factor analysis treat all three modes, and provide results that are sensitive to rearrangements of observations in either space or time. Thus it seems clear to us that such three-way methods are essential to an understanding of space-time behavior. Various three-way models have been described [for a review of three-way methods see Kroonenberg (1983)], and a selection will be reviewed here briefly. INDSCAL (Carroll and Chang 1970) treats one dimension as a replication (Kroonenberg 1983: 54):

$$x_{ii'j} = \sum_p g_{ip} g_{i'p} h_{jp} + U_{ii'j} \tag{3}$$

and extracts one set of factors. The treatment of the third mode as a replication is reflected here by the use of i' rather than k, to indicate a replication of a location rather than a time period. Two matrices of "loadings" are extracted, denoted here by **G** and **H**, and the matrix **U** represents the unique variation present in each observation. Hanham (1976) reanalyzed Taylor and Parkes' data using INDSCAL and reported interesting results, despite the artificiality of the data.

Tucker's model 3 (Tucker 1963; Levin 1963) treats all three modes similarly, by defining a core matrix **C**:

$$x_{ijk} = \sum_p \sum_q \sum_r g_{ip} h_{jq} e_{kr} c_{pqr} + U_{ijk} \,. \tag{4}$$

Three sets of common factors are obtained, one for each mode: p ($p = 1, \ldots, s$) denotes a factor specific to location, q ($q = 1, \ldots, t$) a factor specific to variable, and r ($r = 1, \ldots, u$) a factor specific to time. The matrices **G**, **H**, and **E** describe the relationships between common factors and observed locations, variables, and times, respectively. Finally the elements of **U** define the unique variability present in each observation. The model allows different numbers of underlying dimensions for each mode, and gives them different interpretations, whereas our original hypothesis concerned one set of underlying dimensions appearing in all three modes. Langlois (1983) reported an analysis of the Taylor and Parkes data using this model.

For our analysis we chose the PARAFAC model of Harshman (Harshman 1976; Harshman, Ladefoged, and Goldstein 1977) for two reasons: first, all three modes are treated similarly, which seems appropriate given the nature of the problem; and second, it is the simplest of its class. In the PARAFAC model the contribution of each factor p ($p = 1, \ldots, s$) to the three modes of location, variable, and time is defined by three matrices, denoted here by **G**, **H**, and **E**, respectively:

$$x_{ijk} = \sum_p g_{ip} h_{jp} e_{kp} + U_{ijk}. \tag{5}$$

Thus g_{ip} indicates the degree to which underlying dimension p appears in location i; h_{jp} indicates the same factor's appearance in activity j; and e_{kp} indicates its appearance at time period k. Following common practice in factorial methods, we call the first term on the right-hand side "communality" and the second term "uniqueness."

The PARAFAC model satisfies our basic requirement of modeling variation in all three modes simultaneously in terms of s common underlying dimensions. It makes no assumptions about the relative amounts of each dimension appearing in any time, location, or variable. The one key limiting assumption is that the g, h, and e elements multiply. Thus for a given i, j pair, the contributions of the common dimensions to the observed values of x in two time periods k and k' will always be in the same ratio. As with all such factorial models, equation (5) implies strong assumptions about the form of the common factors' contributions to the observed variables. The goodness of fit of the model provides a measure of the validity of these assumptions.

The data used in this study consist of the numbers of respondents engaged in given activities in given time periods, expressed as percentages of the sample observed to be in a given PCT in a given time period. Because some activities are less frequently observed than others, we normalized the variables across locations and time periods:

$$x'_{ijk} = (x_{ijk} - a_j)/s_j \tag{6}$$

where a_j and s_j are the mean and standard deviation respectively of all x_{ijk} for a given j, that is, the mean and standard deviation of all observed values of a given variable. This normalization provides a better basis for comparison between variables. We also centered the results across locations, yielding a mean result of zero across all PCTs.

The PARAFAC model is calibrated by least squares, that is, by minimizing the sum of squared deviations between the input values x'_{ijk} and the predictions from the common factors, the first term in the right-hand side of equation (5); this is equivalent to minimizing the sum of squared elements of **U**. The sum of squared deviations is expected to decrease monotonically with the number of factors s. As with other factorial models, the advantage of better fit with higher s is offset by the added model complexity and increased difficulty of interpretation. By experimenting with different values of s, we found the solution for $s = 5$ to be the best compromise, and the next section describes the major features of this solution.

TABLE 5
Percentages of Variation Explained by Five-Factor PARAFAC Solution by Time Period

1:00 A.M.–2:00 A.M.	26.2
8:00 A.M.–9:00 A.M.	15.1
9:00 A.M.–10:00 A.M.	37.9
12:00 P.M.–1:00 P.M.	39.2
2:00 P.M.–3:00 P.M.	35.6
5:00 P.M.–6:00 P.M.	24.4
6:00 P.M.–7:00 P.M.	17.8
10:00 P.M.–11:00 P.M.	31.7

INTERPRETATION

Of major interest are the goodness-of-fit statistics by time period, PCT, and variable. Table 5 shows the percentage of variation explained (attributable to common, underlying factors) by time period. Communality is least in the commuting periods, when any part of the city is temporarily occupied by highly diverse respondents, as transition occurs between major activities and associated locations. Interestingly, communality is not as high during the early morning period, dominated by sleeping, as during the working or early evening hours. It appears less easy to explain the geography of nighttime behavior in terms of underlying dimensions, and easiest to explain the geography of midday behavior. We conceive

TABLE 6
Percentages of Variation Explained by Five-Factor PARAFAC Solution by Variable

ACTCOD1	Work income-producing time	65.1
2	Work nonincome-producing time	42.9
3	Travel to and from work	11.7
4	Household work	50.0
5	Household maintenance	11.0
6	Child care	30.5
7	Marketing and shopping for household	71.4
8	Other household work	3.8
9	Travel related to household	18.5
10	Sleep and personal care	49.7
11	Meal times	22.2
12	Educational time	57.9
13	Recreation outdoors and sports	10.1
14	Entertainment away from household	45.3
15	Leisure: socializing	7.0
16	Leisure: radio and TV	9.0
17	Reading, noneducational	6.0
18	Other leisure time	8.9
19	Travel for leisure only	15.0
WHOM0	Alone	20.8
1	With spouse or financé	48.6
2	With child of household	38.1
3	With adult of household	2.6
4	With relative	25.3
5	With colleagues	65.4
6	With organization and club associates	13.9
7	With neighbors	4.7
8	With nonbusiness associates	20.6
9	With others and unknown	22.4
SITE0	At home	85.4
1	At work place	71.1
3	Someone else's building	0.7
4	Outdoors travel by public transit	7.8
5	At public or private building	68.8
6	Leisure indoors	51.1
8	Restaurants and bars	43.7
9	All other places	22.2

of this as a periodic building and breaking down of geographic structure, as the city's population rearranges itself during the morning and evening travel periods.

Table 6 shows the percentage of variation explained by variable. Of the activity variables, marketing and shopping (ACTCOD7) have the greatest communality, as they concentrate strongly in certain time periods and locations. Educational time (ACTCOD12) also shows strong communality, as it also occurs for the 19–64 age group in strongly concentrated locations and time periods. High communalities are also observed for income-producing work (ACTCOD1), household work (ACTCOD4), and sleep and personal care (ACTCOD10), which are also concentrated in space-time.

Of the "whom" variables, activities with colleagues (WHOM5) and spouse or fiancé (WHOM1) show the highest communalities; activities with an adult of the household (WHOM3) are very rare for the respondents (the majority of these are joint activities involving both a 19–64 year old respondent and a parent living in

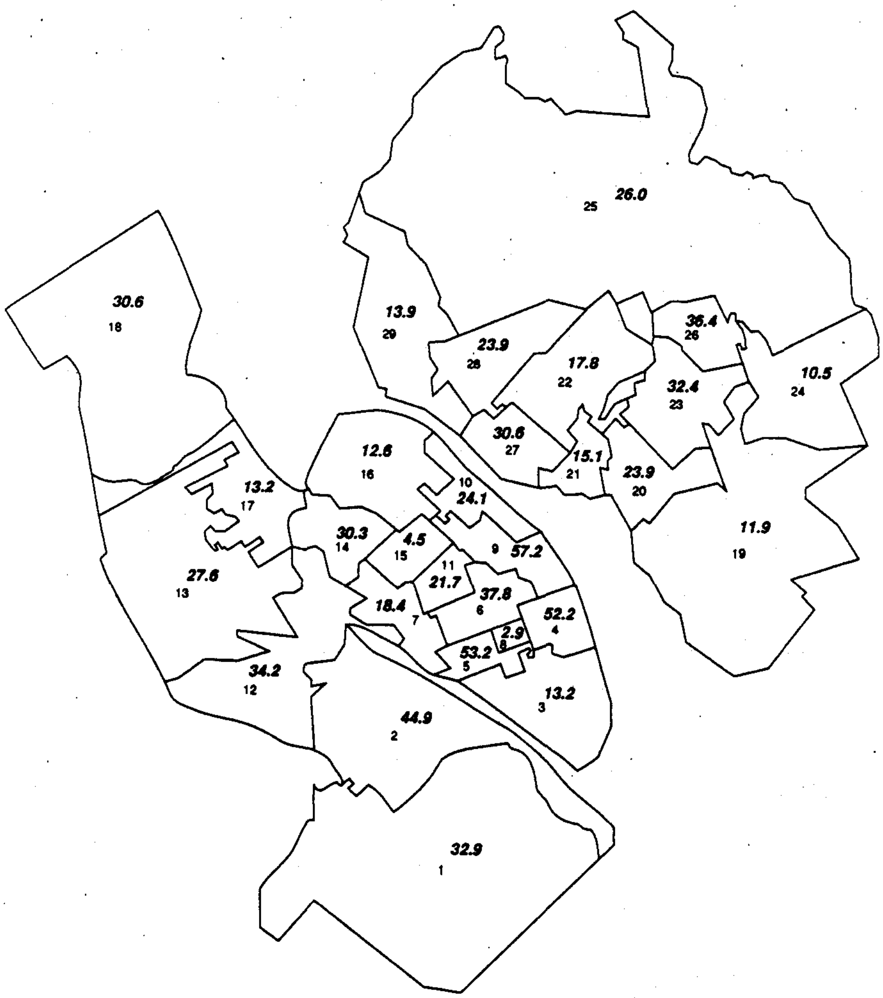

FIG. 2. Percentage of Variation Explained by the Five-Factor PARAFAC Solution in Each of Twenty-nine Pseudo-Census Tracts (PCTs).

the household) and show little space-time communality. The "site" variables show the highest communality for home and workplace.

Figure 2 shows the percentage of variation explained in each of the twenty-nine PCTs. As we might expect from the previous discussion of time periods and variables, the highest communalities are found in the Halifax CBD (PCTs 4, 9), a major concentration of daytime employment; the university area (PCTs 5, 6); the concentration of shopping centers in PCT 14; the area of high daytime employment in the Dartmouth CBD; and in the middle-income suburbs.

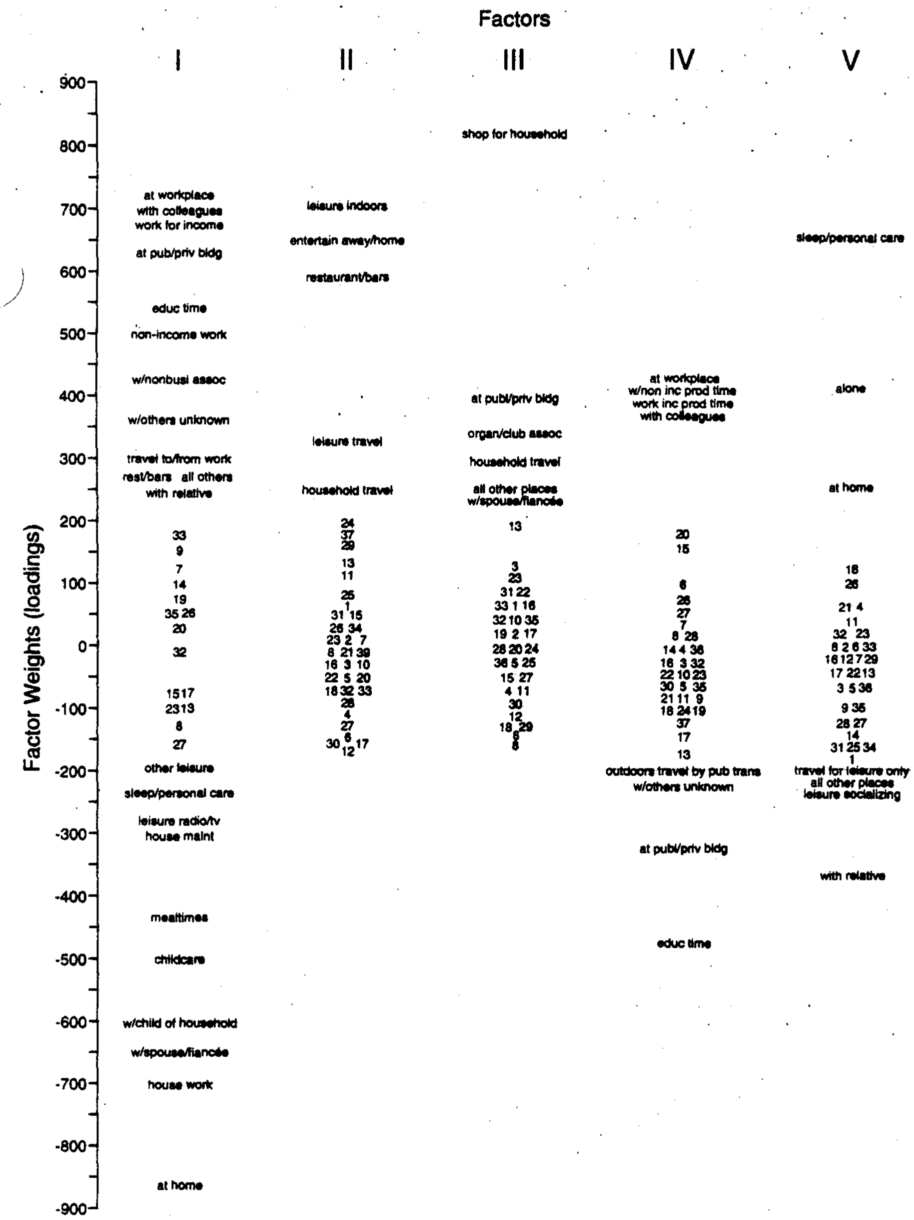

FIG. 3. PARAFAC Solution—Mode A (Variables)

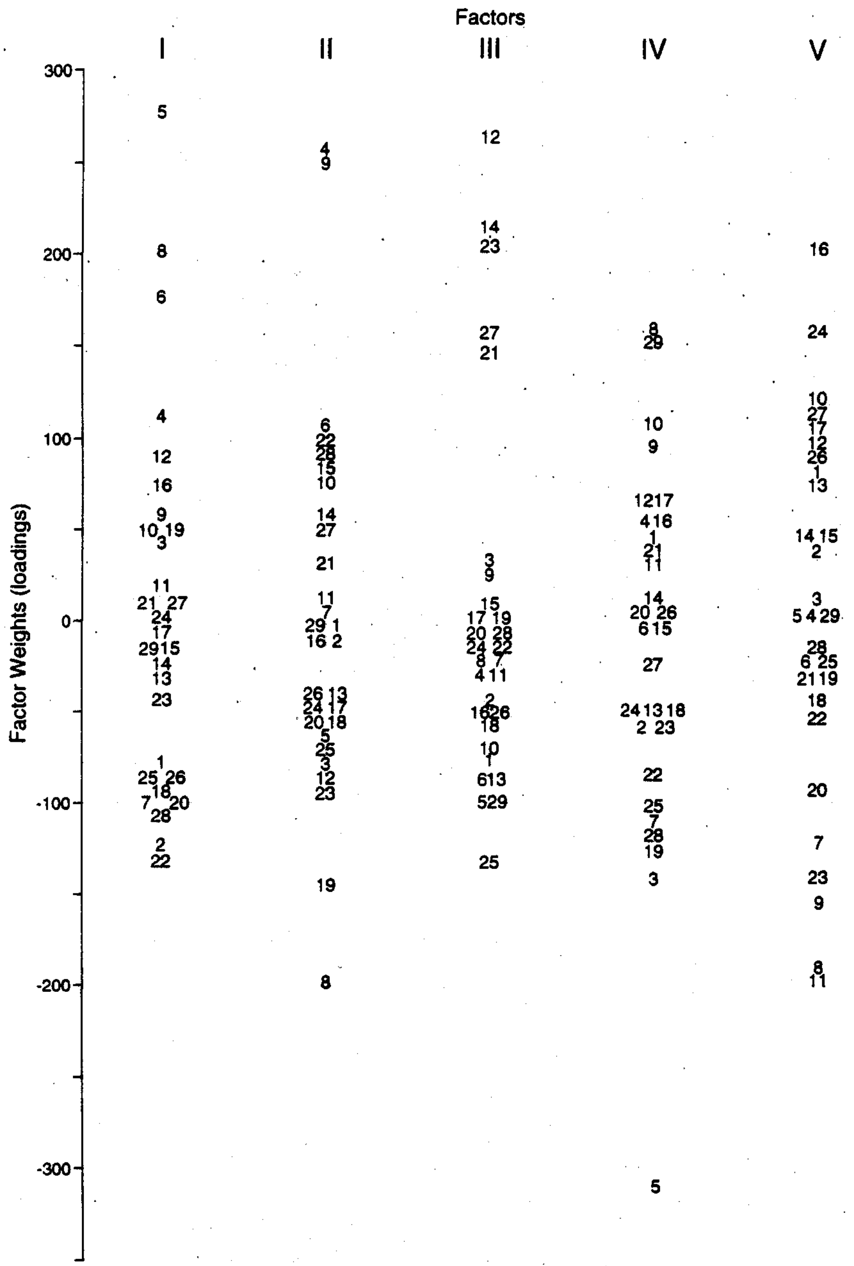

FIG. 4. PARAFAC Solution—Mode B (Space/Tracts)

Figures 3, 4, and 5 show the loadings matrices **H**, **G**, and **E** (variables, locations, and times), respectively, for the five factors. Only the highest and lowest loadings are identified in each case.

Factor I captures the home/work dimension, differentiating between home-related and work-related activities (Figure 3), home and work locations (Figure 4), and home and work time periods (Figure 5). In the **H** matrix dealing with

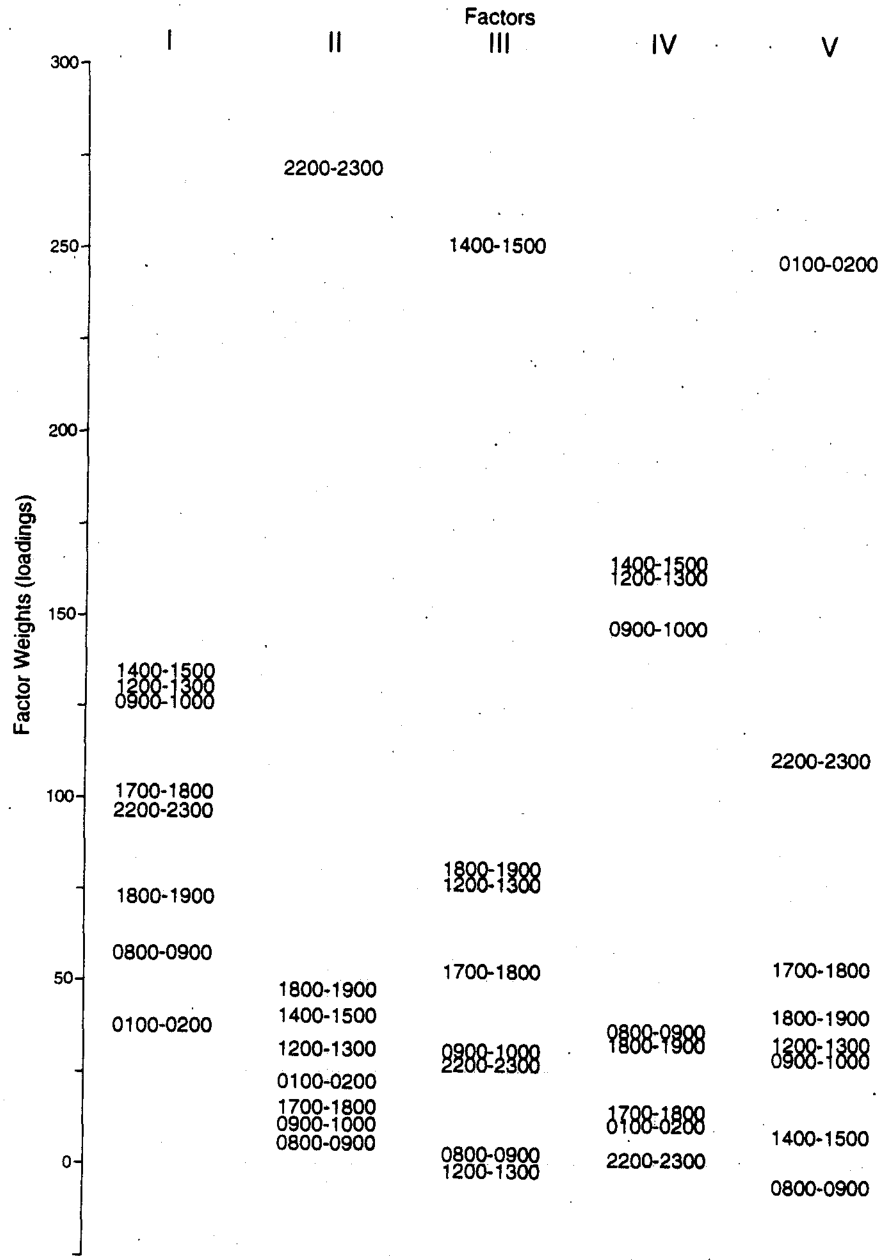

Fig. 5. PARAFAC Solution—Mode C (Time)

variables, the factor is positively associated with "work for income," "non-income work," and "education" among activities, "with colleagues" among the "whom" variables, and "at work place" and "at public or private building" among the "site" variables. Among negative loadings we find "child care," "household work" and "meal time" among activities, "with child" and "with spouse" among the "whom" variables and "at home" among the "site" variables.

In the **G** matrix, dealing with locations, we find Factor I most positively associated with the workplace PCTs, particularly the Halifax CBD and the universities (4, 5, 6, 8, 9) and most negatively associated with the suburban dormitories. In the **E** matrix, relating Factor I to time periods, the most positive loadings are those for 9:00 A.M. through 3:00 P.M. (3, 4, 5) and the most negative for 1:00 A.M.–2:00 A.M., the sleep period. Home/work is the most powerful cleavage revealed by this analysis of the space-time geography of Halifax, and clearly related to activities, locations, and time periods.

Factor II, the second highest in order of contribution to communality, is associated positively with leisure activities ("entertainment away from home" and "in travel for leisure"); and with the "site" variables "leisure indoors" and "at restaurants/bars." The Halifax CBD tracts 4 and 9 contain high concentrations of entertainment facilities, and the highest loading in matrix **E** is in the evening hours (10:00 P.M.–11:00 P.M.). Factor II serves to identify evening entertainment, in terms of activities, times, and locations.

Factors III, IV, and V capture increasingly minor dimensions. Factor III differentiates shopping from other activities, and has its highest loadings in the PCTs with major shopping facilities (12, 14, 23) in the early afternoon period (2:00 P.M.–3:00 P.M.). Factor IV captures the differences between work and education. Although these activities are correlated in space and time, and both contributed positively to Factor I, they are sufficiently independent to be differentiated by Factor IV. Finally Factor V differentiates between solo and group activities, having high positive loadings for "sleep and personal care," "at home," "alone," and in the period 1:00 A.M.–2:00 A.M., and high negative loadings for "leisure: socializing" and "with relative." The high negative loading for the period 8:00 A.M.–9:00 A.M. (Figure 5) may be the effect of time spent taking family members to school or child care.

CONCLUSIONS

The predominantly static view of geographical distributions that is encouraged by the nature of the census of population in most countries, and by current GIS technology, is clearly inadequate for understanding human spatial behavior, and its temporal variability. We have taken the view in this paper that the city is best regarded as an empty space or shell, within which movements occur in diurnal and longer-term cycles. Conventional factorial analyses of cross-sectional census data (for example, Davies 1984) are capable of detecting very long-term changes related to life-cycle status, but tell us little about the massive rearrangements of population that occur in the city every day. By contrast, our study shows the importance of the third dimension in identifying traditionally ignored aspects of human behavior. The census provides one view of the city—one slice of the places/times/variables matrix analyzed in this study. There is a need to incorporate a broader view of the city into a GIS, to consider ways of restructuring the spatial hierarchy that are more reflective of daytime activity patterns.

The ecological characteristics of a neighborhood are a major determinant of human spatial behavior, and are also to some extent a result of that behavior. Thus the socioeconomic characteristics of the residents of a neighborhood affect the likelihood of migrations into and out of the area, and these in turn modify the area's characteristics. Migration affects the spatial structure of the city slowly, on a time scale of years, whereas commuting and entertainment behaviors are determined by similar mechanisms operating on much shorter time scales, of months or even weeks.

While census data are easy to come by, and the census provides some longitudinal coverage of residential locations, there are almost no data readily available on temporal variations over periods of less than five years. The data used in this paper provided individual records of daily activities for a sample of over 1,500 people, but required a series of transformations and discretizations before they could be used effectively in modeling space-time distributions. However, individual diaries represent perhaps the only feasible way of collecting information on space-time behavior. New technologies, particularly cellular phones and the development of "intelligent vehicles," may offer opportunities for studying space-time behavior in the not too distant future, but issues of confidentiality are likely to be important also.

The Census Tract concept used in many countries is an attempt to define longitudinally stable, socially homogeneous areas for analysis. But like the census itself, their homogeneity is entirely defined by nighttime, residential population. The pseudo-Census Tracts or PCTs defined in this paper are an attempt to extend the Census Tract concept to the temporal domain, by measuring homogeneity for the population present in each distinct time period. Residence is only one of a number of significant urban locations occupied by individuals during the diurnal cycle. Patterns of intraurban mobility in the late twentieth century often indicate that place of residence is far from central to daily activities.

Several previous studies have analyzed aggregated space-time data using two-mode methods, by analyzing each time slice separately, or by regarding time slices as additional locations and thus substituting space for time. The results of such methods are invariant under rearrangement of locations within time slices, and thus disregard much of the space-time structure. Three-way factorial methods, such as those discussed in this paper, allow the analysis to be embedded in a space-time framework and thus to test theories of space-time behavior.

The PARAFAC model used in this paper has distinct advantages over other three-mode factorial methods in its simplicity, its homogeneous treatment of the three modes, and its robustness in calibration. Like all models it makes certain assumptions about the form of relationship between underlying factors and observable variables, but these seem parsimonious and intuitively acceptable, and provide reasonable results. By using PARAFAC, we have tested a simple model in which a single set of basic underlying dimensions of space-time behavior appear to different degrees in each time period, location and variable, and in which these contributions combine multiplicatively.

Our diurnal, space-time data set showed strongest differentiation between home- and work-related activities, and identified entertainment and shopping as strong dimensions of the space-time structure as well. The strongest communality, or degree of consistency with basic, underlying dimensions of space-time behavior, was observed at midday, and the least during commuting, and communality is higher during work than during sleep. Marketing and shopping showed the highest communality among activity variables, and on the spatial dimensions, communality was highest in the areas of major daytime concentration of activity. Loosely, we can interpret high communality as indicating simplicity and order, and conclude that the city's space-time structure is most easily understood and least chaotic during working hours and in areas dominated by daytime activity.

Although the data set used in this study is a rich source of information on space-time behavior, it is now over twenty years old, and reveals patterns in only one Canadian city. Unfortunately there are few comparable data sets, particularly ones which provide the same levels of spatial and temporal resolution. An updated Halifax survey might reveal interesting changes in space-time behavior over the past twenty years. We hope that the methods developed in this study and detailed in this paper may provide a framework for comparative studies.

With an understanding of the structure of the city in space and time, it should now be possible to ask questions about the relationship between the behavior of the individual and the structure. How do individuals move within the structure? When and where is individual behavior most strongly linked to the structure, and when and where is it most independent?

Scale is a much neglected factor in the tradition of aggregate urban analysis represented here. The basic mechanism proposed, which links individual behavior to the characteristics of neighborhood, is strongly scale-dependent, but scale is treated only indirectly through the size of the locational unit, whether it be a Census Tract or the PCT used here, or a larger or smaller unit. Unfortunately it is often impossible to examine scale effects because of the nature of the data. In our case, the small sample size of respondents makes it impossible to use many more than twenty-nine reporting zones, but it would be possible to use larger zones, particularly if overlaps were allowed. However, it seems clear that a concerted investigation of scale effects would require a much larger sample of respondents, or a much longer time period, or both.

Similar comments can be made about temporal resolution. The results in this paper were obtained from a discretization of time into eight periods, and while a finer temporal resolution would have been possible, sample sizes would have been correspondingly reduced.

More serious, perhaps, is a general point that can be made about all factorial methods, whether two- or three-mode: that results are invariant under reordering of either the temporal or the spatial dimension. Spatially, this means that any information present in the relative locations of observations is lost to the analysis, and is a serious flaw in a purportedly spatial analysis. It can be solved only by models and methods of analysis that are independent of the space-time sampling frame, and regard space and time as fundamentally continuous (Tobler 1989).

LITERATURE CITED

Berry, B. J. L., ed. (1971). "Comparative Factorial Ecology." *Economic Geography* 47, supplement.

Carroll, J. D., and J. J. Chang (1970). "Analysis of Individual Differences in Multidimensional Scaling via an N-Way Generalization of Eckart-Young Decomposition." *Psychometrika* 35, 283–319.

Clark, W. A. V., M. C. Deurloo, and F. M. Dieleman (1984). "Housing Consumption and Residential Mobility." *Annals, Association of American Geographers* 74, 29–43.

Cosper, R. L, and S. M. Shaw (1984). "The Validity of Time-Budget Studies: A Comparison of Frequency and Diary Data in Halifax, Canada." *Leisure Science* 7, 205–25.

Couclelis, H. (1991). "Requirements for a Planning-Relevant GIS: A Spatial Perspective." *Papers in Regional Science* 70, 9–20.

Davies, W. K. D. (1984). *Factorial Ecology*. Aldershot: Gower.

Elliott, D., A. S. Harvey, and D. Procos (1976). "An Overview of the Halifax Time-Budget Study." *Society and Leisure* 3, 145–59.

Fujita, M. (1989). *Urban Economic Theory: Land Use and City Size*. Cambridge: Cambridge University Press.

Gatrell, A. C. (1983). *Distance and Space: A Geographical Perspective*. Oxford: Clarendon.

Goodchild, M. F. (1988). "Stepping over the Line: Technological Constraints and the New Cartography." *The American Cartographer* 15, 311–20.

Goodchild, M. F., and C. M. Hosage (1983). "On Enumerating All Feasible Solutions to Polygon Aggregation Problems." *Modelling and Simulation* (Proceedings of the Fourteenth Annual Pittsburgh Conference on Modelling and Simulation) 14, 591–95.

Goodchild, M. F., and D. G. Janelle (1984). "The City around the Clock: Space-Time Patterns of Urban Ecological Structure." *Environment and Planning* A 16, 807–20.

Hägerstrand, T. (1970). "What about People in Regional Science?" *Papers of the Regional Science Association* 24, 7–21.

Haggett, P. (1979). *Geography: A Modern Synthesis*. New York: Harper and Row.

Hanham, R. Q. (1976). "Factorial Ecology in Space and Time: An Alternative Method." *Environment and Planning* A 8, 823–28.

Harshman, R. (1976). "PARAFAC: Methods of Three-Way Factor Analysis and Multidimensional Scaling According to the Principle of Proportional Profiles." Ph.D. dissertation, Department of Educational Psychology, University of California, Los Angeles.

Harshman, R., P. Ladefoged, and L. Goldstein (1977). "Factor Analysis of Tongue Shapes." *Journal of the Acoustical Society of America* 62, 693–707.

Hauer, J., H. Timmermans, and N. Wrigley (1989). *Urban Dynamics and Spatial Choice Behavior*. Dordrecht: Kluwer.

Holly, B. P. (1978). "The Problem of Scale in Time-Space Research." In *Timing Space and Spacing Time 3: Time and Regional Dynamics*, edited by T. Carlstein, D. Parkes, and N. Thrift, pp. 5–18. London: Arnold.

Janelle, D. G., and M. F. Goodchild (1983a). "Transportation Indicators of Space-Time Autonomy." *Urban Geography* 4, 317–37.

______ (1983b). "Diurnal Patterns of Social Group Distribution in a Canadian City." *Economic Geography* 59, 403–25.

______ (1987). "The Home-Work Relationship and Urban Ecological Structure." In *Spatial Mobility and Urban Change*, edited by O. Verkoren and J. van Weesep, pp. 39–50. Department of Geography, University of Utrecht.

Janelle, D. G., M. F. Goodchild, and B. Klinkenberg (1988). "Space-Time Diaries and Travel Characteristics for Different Levels of Respondent Aggregation." *Environment and Planning* A 20, 891–906.

Kroonenberg, P. M. (1983). *Three-Mode Principal Components Analysis: Theory and Application*. Leiden: DMSO.

Langlois, A. (1983). "Les transformations de l'espace social de la ville: une application de l'analyse factorielle à trois entrées." *The Canadian Geographer* 27, 67–73.

Langran, G. (1989). "A Review of Temporal Database Research and Its Use in GIS Applications." *International Journal of Geographical Information Systems* 3, 215–32.

______ (1992). *Time in Geographical Information Systems*. London: Taylor and Francis.

Langran, G., and N. R. Chrisman (1988). "A Framework for Temporal Geographical Information." *Cartographica* 25, 1–14.

Levin, J. (1963). *Three-Mode Factor Analysis*. Urbana: University of Illinois Press.

Marble, D. F. (1990). "The Potential Methodological Impact of GIS on the Social Sciences." In *Interpreting Space: GIS and Archaeology*, edited by K. M. S. Allen, S. W. Green, E. B. W. Zubrow. London: Taylor and Francis.

Parkes, D., and N. Thrift (1980). *Times, Spaces, and Places: A Chronogeographic Perspective*. New York: Wiley.

Pickles, A. R., and R. B. Davies (1989). "Inference from Cross-sectional and Longitudinal Data for Dynamic Behavioural Processes." In *Urban Dynamics and Spatial Choice Behaviour*, edited by J. Hauer, H. Timmermans, and N. Wrigley, pp. 81–104. Dordrecht: Kluwer.

Pumain, D. (1991). *Spatial Analysis and Population Dynamics*. J. Libbey Eurotext.

Raper, J., ed. (1989). *Three-Dimensional Application in GIS*. London: Taylor and Francis.

Rossiter, D. J., and R. J. Johnston (1981). "Program GROUP: The Identification of All Possible Solutions to a Constituency-Delimitation Problem." *Environment and Planning* A 13, 231–38.

Sack, R. D. (1980). *Conceptions of Space in Social Thought: A Geographic Perspective*. London: Macmillan.

Taylor, P. J., and D. N. Parkes (1975). "A Kantian View of the City: A Factorial-Ecology Experiment in Space and Time." *Environment and Planning* A 7, 671–88.

Tobler, W. R. (1989). "Frame Independent Spatial Analysis." In *Accuracy of Spatial Databases*, edited by M. F. Goodchild and S. Gopal. London: Taylor and Francis.

Tucker, L. R. (1963). "Implications of Factor Analysis of Three-Way Matrices for Measurement of Change." In *Problems in Measuring Change*, edited by C. W. Harris. University of Wisconsin Press.

van Wissen, L. J. G., and A. Rima (1988). *Modelling Urban Housing Market Dynamics: Evolutionary Patterns of Households and Housing in Amsterdam*. Amsterdam: North Holland.

van Wissen, L. J. G., A. Rima, and P. Nijkamp (1986). "Urban Household Dynamics: Methodology and Application." *Systemi Urbani* 2/3, 179–85.

Wrigley, N. (1988). *Store Choice, Store Location, and Market Analysis*. London: Routledge.