AN APPLICATION OF GIS TO STUDY MIGRATION

Robert N. Martin
Department of Geography
Kutztown University
Kutztown, PA 19530

ABSTRACT Geographic information systems provide a method to evaluate complex geographic phenomena. In this paper an entropy-maximizing gravity model through a GIS is used to replicate the patterns of interregional migration for France using the rail systems. Comparison of actual migrations with estimated migrations indicate a very high correlation.

INTRODUCTION

This paper presents an application of a geographic information system to the study of migration. Specifically, a transportation geographic information system is applied to the interregional migration patterns in France using the rail network to distribute the flows.

Geographic information systems in many ways are the answer many of us prayed for during the "quantitative revolution", when days, weeks and months were spent with an adding machine, hand-held calculator or keypunch machines to enter data, calculate a "statistical" relationship and then even more time was spent mapping the results with traditional cartographic methods. I was reminded of this recently when I mentioned to a colleague that I was using a GIS to map net migration for the counties of the United States and he told me his story of graduate school and working on the same problem for months with pencil and paper.

Over the past several years there have been numerous applications of geographic information systems to problems of land-use planning, site location and natural resource allocation. Most of these applications are predominately map overlay processes as initially proposed by McHarg[1]. These studies have resulted, in many cases, with better allocation of limited resources, a better understanding of the role and impact of spatial distributions and, I believe, a resurgence of geography as a viable field. This paper represents of the application of geographic information systems to the analysis of spatial distributions and the modelling of spatial phenomena.

DATABASE PREPARATION

The geographical database was built from the World Map Series (1:1 000 000) sheets for France using Universal Transverse Mercator, UTM, coordinate systems appropriate to the sheets or portions of sheets, see Table I. A separate file was built for each map sheet. The digitizing was done with the ATLAS*DRAW program from Strategic Mapping under a user-defined coordinate system, i.e., UTM coordinates[2]. Using a 1985 SNCF system map as a guide the rail junctions, rail stations, rail lines, boundaries of regional statistical units, and international boundaries were digitized.

Table I Map Sheets Used

Barcelona, 6th ed., 1959
Bordeaux, 4th ed., 1959
Bruxelles, 7th ed., 1956
Firenze, 7th ed., 1959
London, 4th ed., 1954
Lyon, 7th ed., 1963
Madrid, 5th ed., 1955
Milano, 7th ed., 1959

Editing and File Conversion

The digitized files for each of the map sheets were imported in Atlas*GIS with the projection being the appropriate UTM zone[3]. The files were then projected into latitude and longitude and the eight "sheet files" were merged into a single file for France. Once this file was created, the file was edited. The first step was to edit the rail lines to join the ends of rail segments where the "sheet files" met. The point features representing rail junctions and stations were then coded to indicated the statistical region of the point. Based on this classification of the vertices, each of the rail segments were coded to indicate the origin and destination vertices.

Once the geographic file was edited, an attribute file was then constructed. The attributes associated with the rail lines were mainline classification, electrification, ownership, freight operation, passenger operation and road operation. The attributes were categorized by a nominal classification scheme of zero and one. This classification allows various aspects of the French rail network to be extracted or combined and analyzed as needed.

Transportation Network

The next step in the project was to build the network within TransCAD, a transportation geographic information system, from Caliber Corporation[4]. This was accomplished by extracting the mainline rail network, stations, and statistical region boundaries, and then exporting these files using Atlas Import/Export as Atlas ASCII files[5]. These files were then processed by a MS QuickBasic program written by the author to reformat the Atlas ASCII files into TransCAD ASCII files[6]. Using the Database Build program in TransCAD, appropriate databases for the point file - stations, line file - rail segments, and area file - statistical regions were built.

Once the databases were built, the Network Building procedure in TransCAD was used for the creation of the mainline system, see Figure 1. Based on this network file, the Shortest Path procedure for multiple origins and multiple destinations was used to determine the shortest path based on distance between the twenty-one administrative centers of the statistical regions. Corsica was not included since there were not direct rail links with the mainland. This is the system to which flows would be assigned for actual and estimated

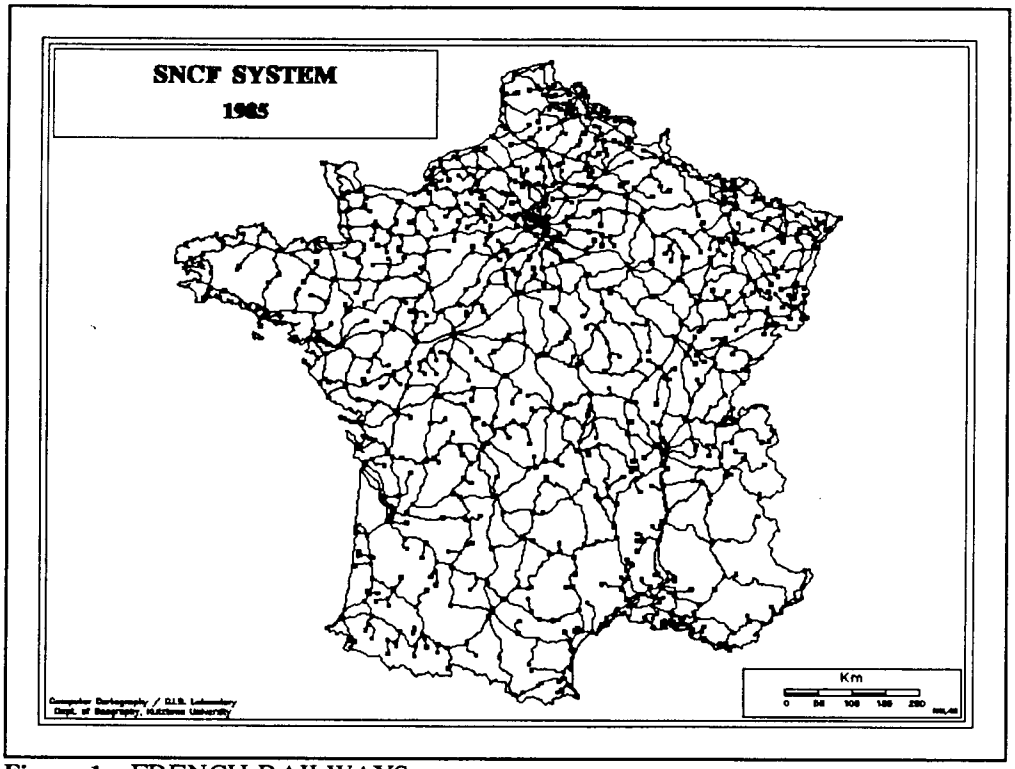

Figure 1 - FRENCH RAILWAYS

migrations, see Figure 2. TransCAD codes the links that represent the shortest paths and creates a distance matrix for the administrative centers.

Migration Data

The final step in the preparation of the database was to build the flow matrix of migrations between the twenty-one administrative regions within the study area. The data represented the average annual migration for 1975 and 1982 between the twenty-one regions based on European Community statistics. No data were available for migrations within the regions and therefore the principle diagonal was left blank. In addition, due to the lack of internal migration information, the distance matrix from the Shortest Path procedure was edited by blanking out the principle diagonal.

The average annual migration between the twenty-one regions was 624,071. Eight of the twenty-one regions had a negative net migrations. These were Ile-de-France, Champagne-Ardenne, Haute-Normandie, Basse-Normandie, Nord-Pas-de-Calais, Lorraine, Alsace and Franche-Comte.

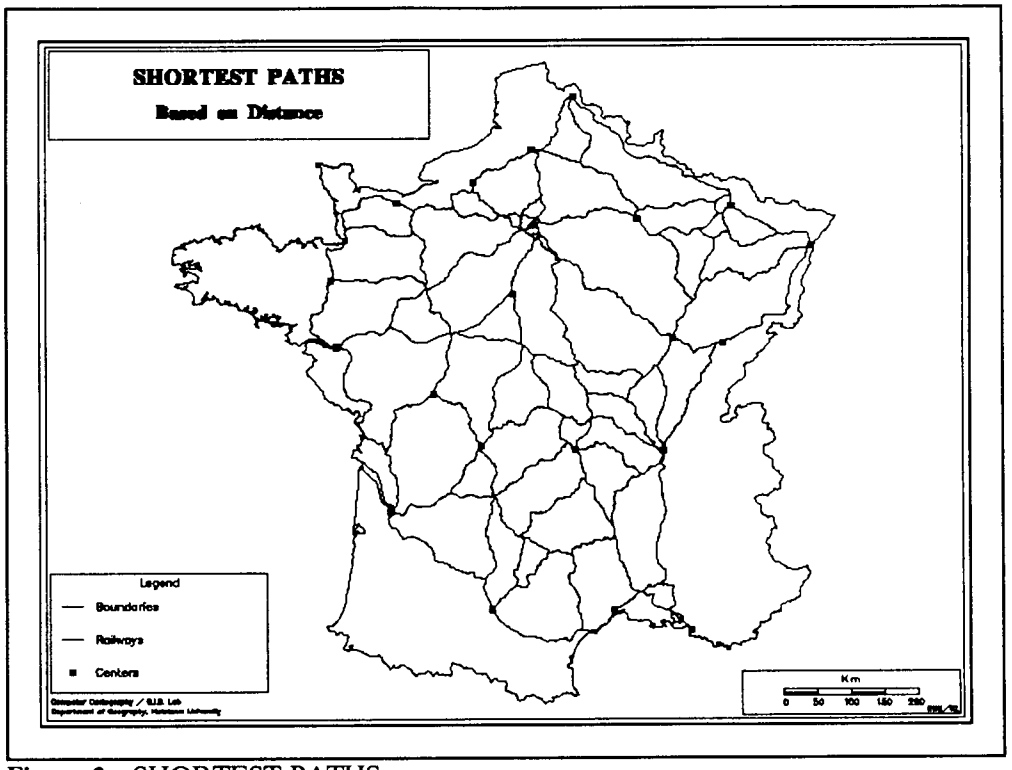

Figure 2 - SHORTEST PATHS

MODEL

A double constrained exponential gravity model was used to model the migration flows. This model was chosen to insure that the number of migrants leaving and entering each region would be correct. The estimated flows were then distributed to the links of the shortest path network as were the actual migration flows by a Traffic Assignment procedure for in-migration and out-migration. A correlation analysis between the estimated flows and actual flows for the origin/destination matrix and in- and out-migration revealed a high correlation, see Table II.

Table II - Correlation Analysis

Analysis	Correlation	No. Obser.	D.F.
Flow Matrix	0.8937	441	439
In-Flow	0.9657	553	551
Out-Flow	0.9257	553	551

As can be seen in Figures 3, 4 and 5 the relationship between the actual migration flows

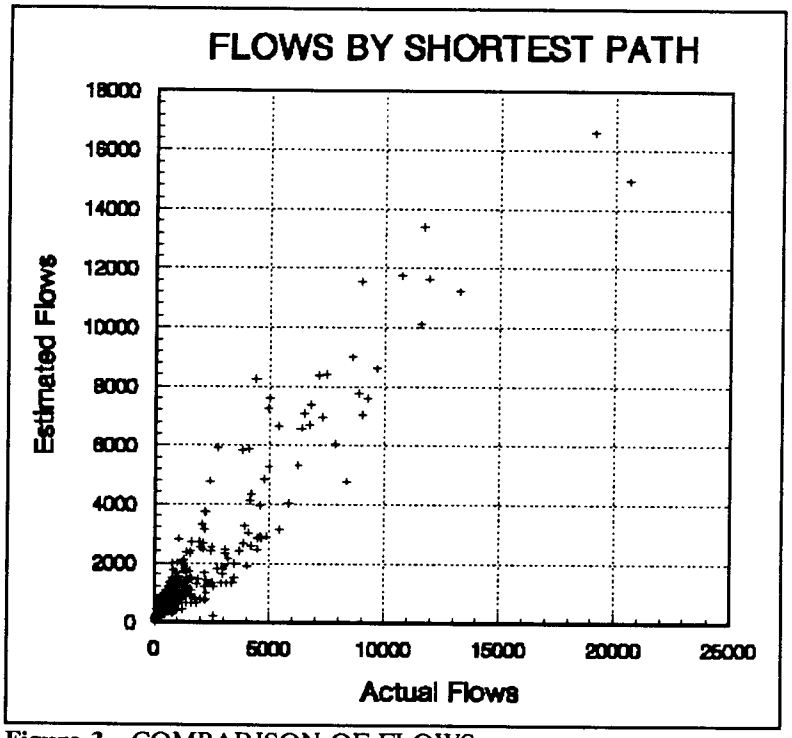

Figure 3 - COMPARISON OF FLOWS

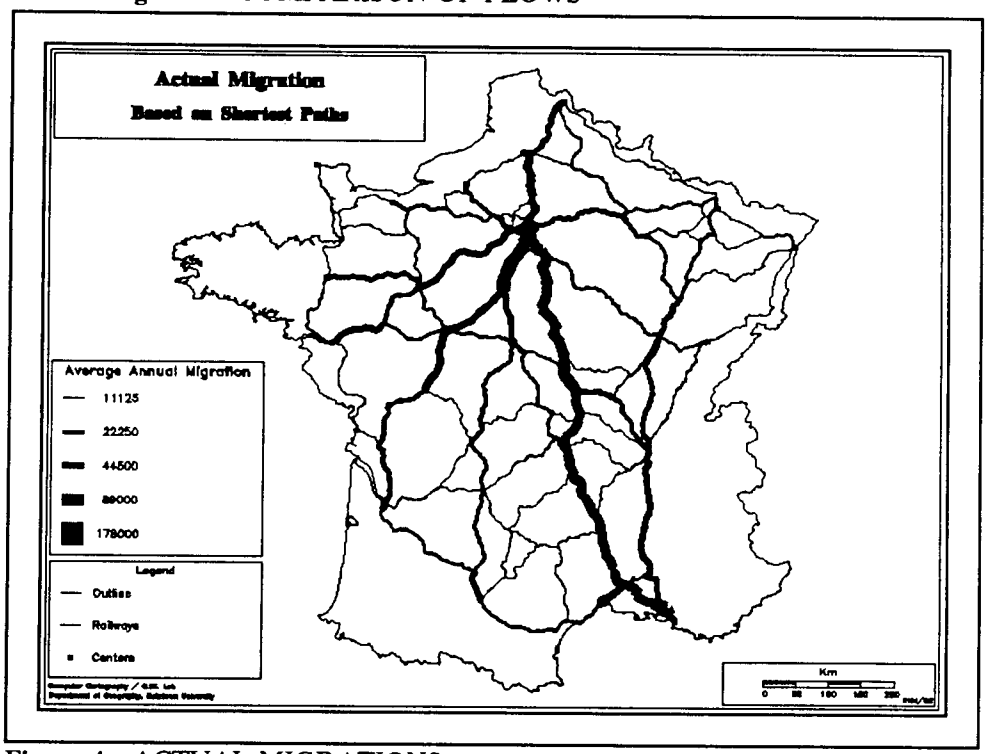

Figure 4 - ACTUAL MIGRATIONS

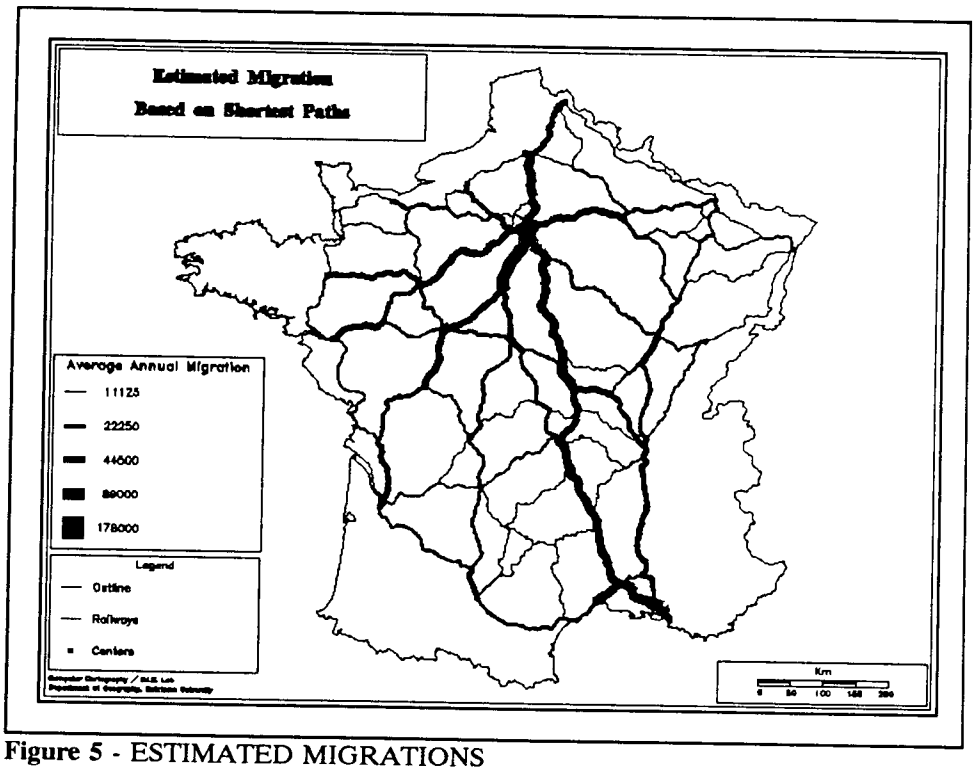

Figure 5 - ESTIMATED MIGRATIONS

and the estimated flows is almost linear. This would indicated that this procedure had produced a very good replication of the migration flows when one considers movements along shortest paths.

REFERENCES

1. Ian McHarg, *Design with nature*. New York: Natural History Press, 1969.

2. *ATLAS*DRAW*, Strategic Mapping, Inc., Santa Clara, CA 95051.

3. *Atlas*GIS*, version 2.0, Strategic Mapping, Inc., Santa Clara, CA 95051.

4. *TransCAD*, version 2.0, Caliber Corporation, Newton, MA 02161.

5. *Atlas Import/Export*, Strategic Mapping, Inc., Santa Clara, CA 95051.

6. *QuickBasic*, version 4.5, Microsoft Corporation, Redmond, WA 98073.