CELLULAR GEOGRAPHY[1]

W. R. TOBLER[2]

Captain Ahab, in the film version of *Moby Dick*, searches for the white whale with the aid of a geographical map on which are noted sighting-frequencies within 5° cells bounded by lines of latitude and longitude. The written version of the story, dating from *circa* 1830, does not contain this scene, but the technique of recording geographical data in this fashion is increasingly popular today. One of the motivations for the use of such partitionings is their 'objectivity'. It is also asserted that there are advantages for analysis purposes over the irregular spatial polygons defined by political jurisdictions. There is no doubt that there are notational simplifications; one can index a cell of an array in the same fashion as in matrix algebra. Thus the cell in the i^{th} row and j^{th} column becomes the cell i,j. Geographical data which pertain to that cell can be referred to by subscripts, as g_{ij} for example. If one lets G represent an N by M array of such cells then this can be considered isomorphic with a portion of the surface of the earth (if one deletes the poles and makes a convention about the edges). But one can also apply matrix algebra to this array and can obtain geographically interesting results. The major advantage however is pedagogical, and results from the fact that in such a scheme every country in the world has exactly the same number of neighbors. The analytical results can be extended to the more realistic variable-number-of-neighbors case but the insight is more easily gained in the cellular case.

I. TYPES OF MODELS

Using the positional notation let g^t_{ij} be the land use category (urban, rural,…) at the location i,j at time t**.** Let $g^{t+\Delta t}_{ij}$ be the land use category at this location at some other time. One primitive classification of models of land use change is then as follows:

(I) The *independent* model: $g^{t+\Delta t}_{ij}$ is a random variable in no way related to g^t_{ij}.

(II) The functionally *dependent* model. The land use at location i,j at time t + Δt depends on the previous land use at that location, $g^{t+\Delta t}_{ij} = F(g^t_{ij})$.

(III) The *historical* model. The land use at position i, j at t + Δt depends on the several previous land uses at that location:

$$g^{t+\Delta t}_{ij} = F(g^t_{ij}, g^{t-\Delta t}_{ij}, g^{t-2\Delta t}_{ij}, \ldots, g^{t-k\Delta t}_{ij}).$$

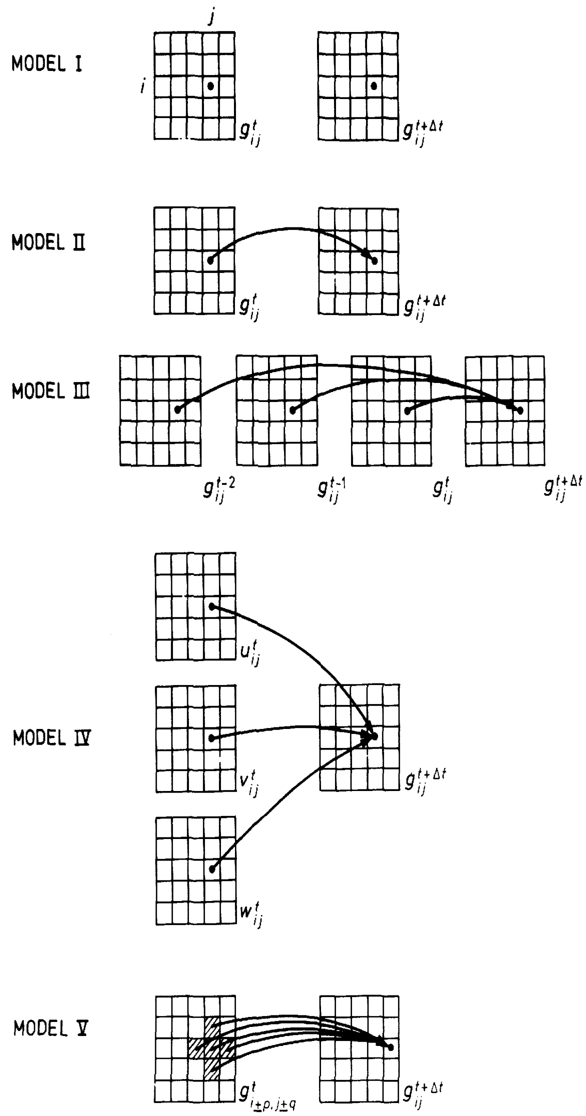
(IV) The *multivariate* model. The land use at location i, j is dependent on several other variables at that location:

$$g^{t+\Delta t}_{ij} = F(u^t_{ij}, v^t_{ij}, w^t_{ij}, \ldots z^t_{ij}).$$

(V) The *geographical* model. The land use at location i, j is dependent on the land use at other (neighboring) locations:

$$g^{t+\Delta t}_{ij} = F(g^t_{i\pm p,\, j\pm q}).$$

Fig. 1. Graphic illustration of the five models using a 25-cell geographical array



These five model types are all simple abstractions from nature. Combining them would be realistic but complicated. One could also embellish these simple types to include moving average models, or stochastic versions - model (II) then readily yields a Markov chain. Model (III) is often called a time series model, or a lagged variable model. Model (IV) could be generalized to a system of simultaneous equations, in which each variable is a function of the several others, and so on. The particular classification has been chosen to highlight the *geographical* model.

There are really two models included in the category of geographical model. The first is the extrapolation-filtering model exemplified by $g^t_{ij} = F(g^t_{i\pm p,\, j\pm q})$. This can be characterized by a geographical quiz:

Complete the following geographical sentence by filling in the blank cell:

A	A	A	A	B
A	A	A	B	B
A	A		A	A
A	B	B	B	A
B	B	A	A	A

There is considerable literature on this topic, but the model of concern here is the dynamical geographical one which is better characterized as

$$g^{t+\Delta t}_{ij} = F (g^t_{ij} , n_{ij}),$$

where n_{ij} is shorthand for all of the land uses in the neighborhood of the location i, j. This single lag, univariate deterministic - as here described - model has only two parameters: the neighborhood n and the function F.

II. NEIGHBORHOODS

The simplest definition of a neighborhood in a square lattice is to include all cells in a box around the cell of interest; n_{ij} = cells $i \pm p$, $j \pm q$. The neighborhood then consists of $(2p + l)(2q + 1)$ cells. Also common is the five cell neighborhood consisting of a cell and its North, South, East, and West adjacent cells.

The importance of the neighborhood is that it defines the geographical domain of influence. But the definition of the neighborhood of a cell can be quite general. One could, for example, provide a list of all of the cells that are included in the neighborhood of a given cell. But the usual rule is to invoke *spatial neighborhood stationarity.* By this is meant that all cells have the same size and shape of neighborhood. The indexing by subscripts, $n_{ij} = g_{i \pm p,\, j \pm q}$ makes this very clear.

This model contrasts very nicely with reality in which, for example, an urban resident may have a geographical contact field that differs in size and shape from that of a rural resident, or of a suburbanite. Thus it is possible to let the size, shape, or orientation of a neighborhood be a function of the location of the cell, i.e., p, q = F(i, j) in either a simple or a complicated fashion; neighborhoods near borders usually require a special definition.

Board games such as chess, checkers, and go are all defined on square lattices; Chinese checkers on a triangular lattice. One can see the advantages of such arrays most easily if one attempts to define a game similar to chess on a political map. An identical problem is encountered in converting geographical lattice models - Hägerstrand's model of the diffusion of ideas, for example, to political units. The basic difficulty is topological; the 'cells' on the political map do not all have the same number of adjacent cells. Their neighborhoods cannot be defined by any simple notational scheme, and the concept of spatial stationarity of neighborhood must be defined in a different manner.

III.THE TRANSITION RULE

The other important variable in the geographical model $g^{t+\Delta t}_{ij} = F(g^{t}_{ij}, n_{ij})$, is the function F. For the present purpose it is still valid to ignore such distinctions as deterministic or stochastic, time varying, and so on, and to concentrate on the geographically interesting aspects. An example is helpful. Suppose that the contents of the geographical cells consist of five land use types: Residential (R), Commercial (C), Industrial (I), Public (P), and Agriculture (A). Suppose further that the neighborhood consists of the cells (i,j), (i-1,j), (i+l,j), (i,j-1), and (i,j+1). There are thus five states (S) and five neighbors (N). A possible transition rule would be

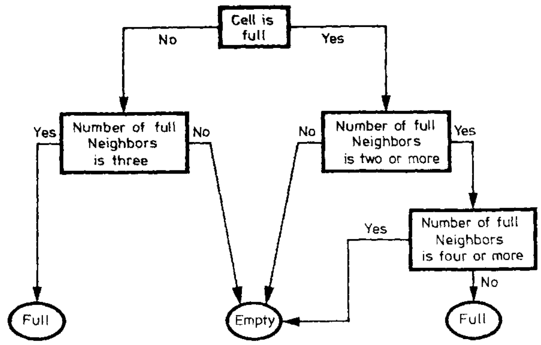
$$\begin{matrix} & R & \\ R & A & I \\ & C & \end{matrix} \rightarrow \begin{matrix} & R & \\ R & C & I \\ & C & \end{matrix}$$

which means that the center cell, in agriculture, is converted to a commercial land use. This might more conveniently be written as RICRA → C with a clockwise convention. One sees that one must consider the $S^N = 5^5 = 3125$ cases to cover all possibilities. But it is now natural to invoke *spatial isotropy* so that the positioning of the neighbors does not count, e.g., writing the above rule as (2R, lI, 1C, A) → C and this clearly cuts down on the necessary number of rules. Of course we have already assumed *spatial stationarity* again. Translated this means that the same environment (neighborhood) results in the same consequences, or, that the rules do not depend on where you are. Compare again with chess; the allowed moves, although piece specific, are the same everywhere, almost. Thus, the laws of nature do not depend on, say, latitude. Or do they? ‘When in Rome do as the Romans do’. Cultural geographers assert that behaviour in England is different from that in China. This is equivalent to saying that the rules depend on where you are, i.e., F(i, j). These models and games make a nice pedagogical contrast with reality. Sometimes it is easy to write down rules that depend on where you are, sometimes it is not easy.

One type of scientific investigation can be caricatured by the following problem: given 20 pictures, in order, of the board positions from a game of chess, determine the rules of chess. The rules of chess are rather simple, but the game, which involves using the rules in a strategy, is complex. Does a similar situation hold for changes of pattern on the surface of the earth? My students have now conducted some experiments in which geographical maps (of one area but from different times) are fed into a computer and a program attempts to estimate the geographical transition rules.

An analogy can also be made with geographical planning. Given an initial state, a desired state, and a set of transition rules, we can ask whether or not there exists a path from the one situation to the other, and if so, whether there is a minimum path. Or what changes need be made to the rules so that the objective is realizable.

Some of these ideas are nicely illustrated by Conway’s ‘Life’, a two-state, nine-neighbor play. The game is played on a square lattice, and the two states are conveniently called filled or empty.



The change of state from full to empty, or visa versa takes place via the rules, which are conveniently displayed as a decision tree, invoked for all cells simultaneously in one round of the play.

The play begins from an initial state in which some cells are full and others are empty. Such a pattern then changes over time, appearing to move, often in interesting ways. Sometimes the pattern repeats itself periodically, in other cases it disappears completely. One can prove *inter alia* that there exist patterns which could never arise from some other initial state (Moore's Garden-of-Eden theorem). The point that I wish to stress is that there are a whole host of theoretical questions that one can ask of even such a simple two-state nine-neighbor situation, and similar theoretical questions should also be asked of the more complex geographical case.

As a final contrast attention is called to the fact that all of the examples considered up to this point have been of the categorical type. On rare occasions in nature the states, the observed entities in the cells, can be represented by numbers. This rare situation seems to be the one most often studied by scientists in general and geographers in particular. The transition rule in this case becomes the usual mathematical function. Of all possible functions linear functions are most often used, e.g.,

$$g^{t+\Delta t}_{ij} = \sum_{-p}^{+p}\sum_{-q}^{+q} W_{pq} g^{t}_{i+p,j+q}\,,$$

in the discrete case, or

$$g^{t+\Delta t}_{ij} = \int\int_{-\infty}^{+\infty} w(x-u, y-v) g^{t}(x, y) du dv,$$

in the spatially continuous case. Here one again notices the spatial stationarity assumption. Possible geographical interpretations and applications have been discussed elsewhere and need not be repeated here. Perhaps more results can be expected from a study of the above case, but an interesting area would appear to be in the mathematical study of non-numerical transformations.

NOTES

A condensed translation of a lecture "Schachbrett Modelle in der Geographie" presented to the Arbeitskreis für neue Methoden in der Regionalforschung, Wien, 17 April 1975.

[1]From: S. Gale and G. Olsson (eds.), *Philosophy in Geography*, 379-386.

[2]Department of Geography

University of California
Santa Barbara, CA 93104-4060, USA.

BIBLIOGRAPHY

Agterberg, F., 1974, *Geomathematics,* Elsevier, New York.

Codd, E., 1968, *Cellular Automata,* Academic Press, New York.

Hägerstrand, T., 1968, 'A Monte Carlo Approach to Diffusion', in B. Berry and D. Marble (eds.), *Spatial Analysis,* Prentice Hall, New York.

Matheron, G., 1971, 'The Theory of Regionalized Variables and Its Applications', *Cahiers* du Centre de Morphologic Mathematique de Fontainebleau, No. 5, Ecole National Superieure des Mines.

Tobler, W., 1975, 'Linear Operators Applied to Areal Data', in .J. Davis and M. McCullaugh (eds.), *Display and Analysis of Spatial Data,* J. Wiley, New York.