A mathematical approach to urban design--a progress report on a land use plan design model and a land use simulation model. no. 3 January 1966

[s.l.]: Southeastern Wisconsin Regional Planning Commission, January 1966

https://digital.library.wisc.edu/1711.dl/VRTZAXYXLJUWG8K

TECHNICAL REPORT

NUMBER 3

A MATHEMATICAL APPROACH TO URBAN DESIGN – A PROGRESS REPORT ON A LAND USE PLAN DESIGN MODEL AND A LAND USE SIMULATION MODEL.

Prepared as part of the
Southeastern Wisconsin Regional Planning Commission
Land Use-Transportation Study

The preparation of this publication was financed in part through a joint planning grant from the State Highway Commission of Wisconsin; the U. S. Department of Commerce, Bureau of Public Roads; and the Housing and Home Finance Agency, under the provisions of the Federal Aid Highway Legislation and Section 701 of the Housing Act of 1954, as amended.

JANUARY 1966

Inside Region $1.50
Outside Region $3.00

Chapter I

INTRODUCTION

THE FUNCTION OF MATHEMATICAL MODELS IN LAND USE PLAN DESIGN AND IMPLEMENTATION

Two Basic Problems: Plan Design and Plan Implementation

To be successful in the attempt to improve the urban and rural environment, land use planning must deal with two complex interrelated problems. The first problem is that of plan design which relates to the synthesis of a plan for the future spatial location of land use activities that satisfies some set of plan objectives. A second problem, for which the first is a necessary prerequisite, involves the implementation or realization of the previously designed plan in the real world.

This report explains the application of applied mathematics and electronic computation in the form of a pair of mathematical models to the dual problems of land use plan design and implementation. The Land Use Plan Design Model is intended to deal better with the first of these problems. The Land Use Simulation Model is intended to deal with the second of these problems. The conceptual bases, theoretical development, and practical application of these two models are the subject of this report.

The complexity of these two problems and the inherent limitations of current conventional land-use planning methods to deal with this complexity justify the expensive and time-consuming effort required to develop and apply mathematical models to the land-use planning problem. Although few individuals would now question the complexity or importance of these two planning problems, some would suggest caution in ambitious claims regarding their impending solution. The complexity of these problems may be so abstruse and elusive as to defy the present state of the art in both mathematics and computation. Such a reservation would be directly related to two factors underlying the current decline of the design approach to land use planning:[1]

> First, as a result of technological advance and social change, the size and complexity of our urban concentrations has grown enormously. Their function and growth patterns now surpass the intuitive understanding and powers of normative reduction of any single individual. Second, the relative expansion of the private market economy in urban land, and the growth of a pluralistic society, have greatly complicated the processes of decision-making and control in urban development. The master-builder can no longer impose his will upon all groups and individuals who by their actions, contribute to change and to the emerging pattern of urban form at any particular time.

The first factor in the decline is really a critique of design technology and applies to the whole field of design, not just to urban design. Despite the rapid advance of science in recent decades, the technology of design has lagged significantly. The design process has become extremely complex and design engineers not only have a much wider area of choice of materials and processes available for new products but must design these products to meet an increasingly complex set of requirements. Techniques of design have not advanced to keep pace with the needs for making better design choices in an increasingly complex situation. In some cases, the wider range of choice actually has been detrimental to the final quality of the product, as evidenced by the fantastic rate of failure of new products in industry.

What is true for product design is accentuated for urban plan design. The size and complexity of urban form and the dynamic nature of urban development have indeed reached incredible

[1] *Britton Harris, "Some Problems in the Theory of Intra-Urban Location"; Penn-Jersey Transportation Study, P. J. Paper No. 3, 1962.*

proportions in modern times, and some radically new innovations must be developed to cope with the completely changed nature of the urban design problem. It is hard to contest the inability of any one designer, however great, to intuitively manipulate all the variables involved in a complete urban plan design.

With the aid of mathematical models, however, and with the means of practical application of such models available in the computer, the many variables and conflicts involved in urban design may not only be amenable to resolution but to resolution in an optimal fashion in accord with plan objectives.

The second factor in the decline of the design approach to land use planning implies that, even if urban complexity can be overcome, the plan design cannot be implemented because of the pluralistic distribution of decision-making in our society. Difficulties encountered in urban plan implementation support this assertion. The situation is not hopeless; however, the problem must first be clearly and specifically stated. One such way of stating the problem is as follows:

1) The objective of land use plan implementation is to bring about the target plan design. The variables describing this plan (land use activities in different areas) are known as the target variables.

2) These target variables may be influenced by certain other variables known as the controlled variables which are subject to governmental decisions. These variables are those associated with public works programming and public land use controls.

3) These target variables are also influenced, however, by uncontrolled variables determined by private decision-makers, such as land developers, builders, and households.

The problem of land use plan implementation then, stated succinctly, is how to achieve a given set of target variables representing the land use plan design using the controlled variables and considering the possible adverse influence of the uncontrolled variables. The situation resembles that of a ship captain piloting his vessel toward home port (target variable), keeping the ship on the required base course through utilization of the rudder and engines (controlled variables), and in the presence of wind and seas (uncontrolled variables) continuously driving the ship off its course. The purpose of the Land Use Simulation Model is to help determine the steering and engine room signals (controlled variables) needed to guide land development so as to reach the home port of the plan design (target variables).

LAND USE-TRANSPORTATION PLANNING PROCESS

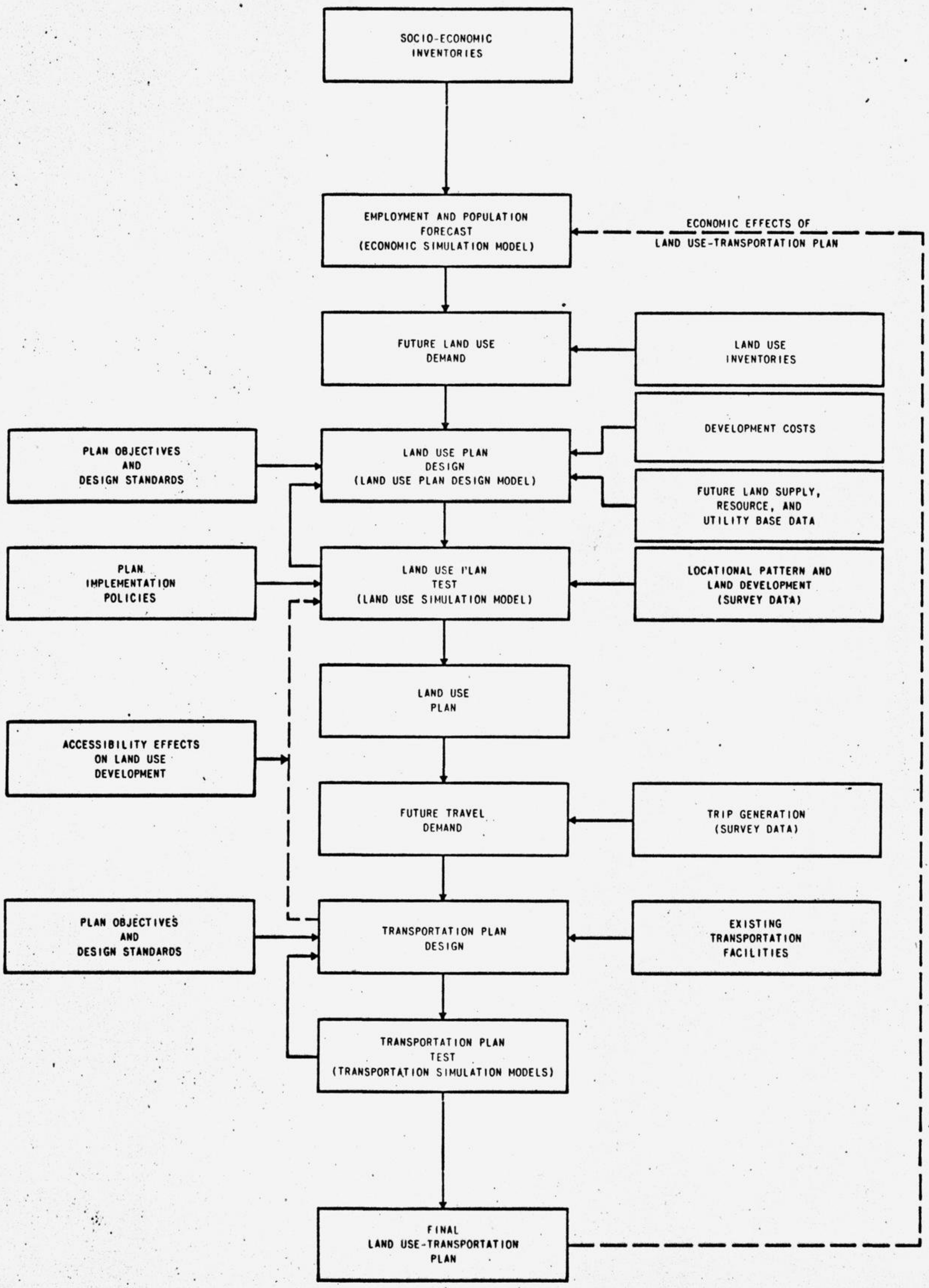
The subjects of this report are land use plan design and implementation and the related land use plan design and simulation models. Land use plan design and implementation, however, are only two of a sequence of functions in the planning process. For this reason, the introduction of the models will be preceded by a brief discussion of the role of mathematical models in a specific land use-transportation planning sequence.

A system diagram illustrating the functional relationships in the land use-transportation planning process is shown in Figure 1. Although this diagram specifically represents the planning sequence related to the formulation of a regional land use-transportation plan, it is typical of other land use facility planning sequences.

The first function in the planning sequence is that of population and employment forecasting. Because population and employment are the primary determinants of land use requirements, they must be forecast as a preliminary to the determination of future land use requirements. In the land use-transportation planning program of the Southeastern Wisconsin Regional Planning Commission, new methods of socio-economic forecasting are being investigated in an

Figure I

LAND USE TRANSPORTATION STUDY PLANNING SYSTEM DIAGRAM



attempt to provide more accurate and comprehensive employment and population forecasts. These new techniques, which center around a regional economic simulation model, are the subject of a companion technical report [2] and will not be discussed in detail here. Whatever the method used, population and employment forecasts must be provided as the output of the first step of the planning sequence.

In the second function, aggregate land use demand requirements are determined by applying a conversion coefficient, usually designated as a design standard, to each employment and population category. Such a multiplication and summation will result in a detailed classified set of aggregate demands for residential, industrial, commercial, and other land uses. These aggregate demands provide one of the primary inputs to the third function, plan design.

Plan design lies at the heart of the planning process. The land use plan design function consists essentially of the allocation of a scarce resource—land—between competing and often conflicting land use activities. This allocation must be accomplished so as to satisfy the aggregate needs for each land use and comply with all of the design standards derived from the plan objectives at an acceptable cost.

The plan selected in the design stage of the planning process must be implemented in the real world under conditions often adverse to its realization. Private decisions of land developers, builders, and households may run contrary to the development of the land pattern prescribed in the plan. This problem of plan implementation is the function of the third stage of the planning process, illustrated in Figure 1, land use plan implementation test.

Land use plan implementation is simulated in the Land Use Simulation Model by detailed representation of the decision processes of households and business firms influential in land development. Public land use control policies and public works programs are exogenous inputs to the model. In practice, a number of experimental simulation runs must be performed with different land use control policies and public works programs until a set of policies and programs are determined that result in the implementation of the target land use plan. The feedback on the diagram between land use development and land use plan design accounts for changes that may have to be made in the plan design to make it attainable. The output of the third stage of the process is a land use plan capable of practical implementation.

The remaining stages of the planning sequence depicted in Figure 1 relate to the development of a transportation plan. The primary inputs to a transportation system are the trips generated as a function of land use. For this reason, the land use plan is shown in the diagram as an input to the transportation plan design. It will be noted that no models are indicated in the transportation plan design function. None are known to exist at this time. Trip distribution, modal split, and traffic assignment models may be used to test the plan intuitively designed by the transportation planner. As a result of model simulation, the transportation plan network is revised until a satisfactory system is developed.

Although each function in the planning process is important to the final realization of a creative and practical plan, the vital role is played by plan design since it is the focal point of all preceding and succeeding planning activity.

DESIGN AND SIMULATION MODEL CHARACTERISTICS

The Land Use Plan Design Model and the Land Use Simulation Model represent two contrasting types of mathematical models. The Land Use Plan Design Model is a design, or by the more common designation, a normative model that provides an efficient search procedure for evaluating alternative plans in the light of stated objectives and within the boundaries of related constraints. In short, this design model specifies a desired land use pattern.

[2] *SEWRPC Technical Report No. 5, The Regional Economic Simulation Model, Theory, and Application, 1965.*

Actual and not necessarily desired land use development is simulated by the Land Use Simulation Model. This simulation model is a positivistic model in that it represents the actual rather than the ideal behavior of private land development decision-makers. Land development is represented as a dynamic process which is continually changing over time.

Both of these models may be more fully understood through a clarification of the nature of models and particularly their role in architecture and engineering. Models are representations of the real world that are used to explain or modify some aspect of it. Originally, models were confined to physical representations of structures. These physical models were related in scale to the objects represented. The basic purpose of these physical models was that of dimensional analysis. The dimensional relationships of a building or bridge could be analyzed using the model since there was a geometric similarity between the model and the real object.

At first, the dimensions of physical models were restricted to static characteristics, such as height, width, length, and weight. Later, the dynamic dimension of time was added; and aircraft and ship models not only "looked like" their real world counterparts but "acted like" them over a period of time. The dynamic characteristics of ships and aircraft could be determined in model water basins and wind tunnels just as the static dimensional characteristics of bridges and buildings had been.

Architects and engineers using these physical models as aids in design understood their limitations. They realized that they were only approximations of the real object in question. Certain detailed characteristics of the object were not well represented in the model. The key question, however, related to the degree of accuracy of the approximation. If the representation was accurate enough for the relationship being studied, it was useful in description or design.

A mathematical model is like a physical model in that it is a representation of the real world. Instead of physical dimensions, however, the similarity is expressed in mathematical symbols. Physical dimensions of the modeled object are represented as algebraic variables in a mathematical relationship such as an equality. A very simple example of such a static model is the equation which represents the weight of a cube by the third power of its side dimension multiplied by the density of the material involved. This mathematical model resembles the real object in the sense that it embodies the same relationships between weight, side dimension, and density as the real object.

A mathematical model becomes dynamic when time is included as one of the variables. A very simple example of such a dynamic model is the equation which represents the velocity of a falling body accelerated by gravity as a function of time. A slightly more complex such model represents the vertical motion of an oscillating weight on a spring. Dynamic models differ from static models in the interaction of other variables with the variable of time.

Because of the logical relationship between mathematics and language, a mathematical relationship is equivalent to a sentence in language. The cube and vehicular velocity models discussed above could be expressed in words, as well as mathematical symbols, as in fact was actually done in the previous paragraphs. A language, such as English, however, does not provide an ideal vehicle for the expression of models because of the multiple meanings of words and the difficulty in manipulating verbal statements. Mathematics allows for more precise definition of variables and facilitates the manipulation of complex relationships.

The language analogy does assist, however, in clarifying the function of a mathematical model as a statement of an outcome. The cubic weight model states that the outcome of a given side dimension and material density will be a given weight of the cube. The dynamic model of a falling body states the velocity "outcome" of a falling body after a given period of time.

Most models are statements of the value, outcome, or output of one dependent variable depending on the values of other independent variables. A special class of models, important in planning, are optimal-value or normative models which are capable of determining the values of the variables that result in an optimal outcome. The Land Use Plan Design Model is an example of such a model. This design model optimizes the cost variable by determining the land use plan that minimizes development costs while complying with land use demands and design standards.

The second land use model, the Land Use Simulation Model, is not an optimal-value model. It is only an outcome-producing model which includes the variable of time. Such models are generally designated as simulation models because they simulate the dynamic sequence of some process. Strictly speaking, simulation models would not need to include the variable of time; and in this sense all models, other than optimal-value models, are simulation models. A simulation model that includes the time variable is more strictly a dynamic simulation model.

It is important to realize that the concept of a model is not entirely new. As has been previously pointed out, physical models have been used by architects and engineers for some time. Moreover, since a model, broadly defined, is a representation of reality, then all thought is a model since it is a representation of reality.

It is also necessary, however, to understand what is new in the concept of simulation models and their recent application. Optimal-value models, while not strictly new since the calculus was used to determine maxima and minima, have reached a high state of development in the last two decades. Mathematical programming in all of its forms—linear, nonlinear, and dynamic—has made the application of optimal-value models a practical reality. Prior to the development of the Simplex Algorithm by Dantzig after World War II, linear programming in most of its current applications was not practical. A model, such as the Land Use Plan Design Model, even with the largest and fastest digital computer, could not be developed without the techniques of linear and dynamic programming.

Since the objective of science is to describe nature, optimal-value models have little use in this area (unless nature is believed, as in some economic-ecological theories, to act in an optimal manner). In the applied fields of business management, architectural and engineering design, and urban planning, however, optimal-value models are at the core of the task to be accomplished, which is to apply nature to the purposes of man.

Dynamic simulation (process) models, in a generic sense, are less new than optimal-value models since any equation that includes the variable of time is, in a sense, a dynamic simulation model. What is new is the practical reality of large-scale simulation models. Prior to the advent of the digital computer, large-scale simulation models could be formulated but not economically solved. One solution of the model could consume many man-years of human effort. With an electronic digital computer, however, large-scale simulation models of the type represented by the Land Use Simulation Model may be computed in a reasonable time and, therefore, become practical tools for planning application.

Both the regional economic simulation and the land use simulation models, developed by the Southeastern Wisconsin Regional Planning Commission, are dynamic process models which generate a synthetic history of the system variables over a period of time. Starting from a given set of initial conditions, the difference equations used in the model permit the calculation of the change in the system variables during the first time interval. The new state of the system then becomes the new base for the change computations of the second time period. If A is the initial residential land area and a function dR expresses the rate of change in residential land use in a given time period, then:

$$R_t = R_{t-1}+(dT)(dR)$$

where

$$R_o = A$$

and

$$dR = f(x_1, x_2 --- x_n)$$

R_t - Residential land area

dT - Recursive time interval

dR - Rate of change of residential land use

$x_1, x_2 --- x_n$ - Other model variables influencing the rate of change of residential land use

In general, the difference equations are sequential[3] rather than simultaneous, although an exception to this general rule exists in the Land Use Simulation Model which has both simultaneous and sequential relationships.

Both the regional economic and land use simulation models are made up of a large number of equations of the type illustrated above. Four classes of problems[4] exist in the development of simulation models of this kind: 1) the formulation of the basic functional relationships involved in the model, 2) the development of a computer program of the model, 3) the estimation of the parameters for the model relationships, and 4) the validation of the model.

The rationale for each of these problems in the Land Use Simulation Model will be explained and related to the current state of model development.

It is well to recognize that these simulation models represent only a part of a large number of similar model development efforts in urban planning and other fields now underway in this country and other parts of the world. The work of the Soçial System Research Institute at the University of Wisconsin on national economic simulation models,[5] the program of Jay Forrester and his associates in industrial dynamics at the Massachusetts Institute of Technology,[6] and many unpublished proprietary simulation models developed by individual industrial firms are only a few of the programs, proceeding along the same general lines.

Most land use models of the non-design variety, aimed at forecasts of future land use, however, have not been dynamic simulation models but, rather, single-stage forecasts of land use for a given point in time. An exception to this general situation is the Penn-Jersey regional growth model, which combined simulation and linear programming using a five-year iteration time. Model practitioners have generally recognized the inherent desirability of a dynamic simulation model, but most projects have been limited by a lack of data and have had to make use of data collected for other purposes.

---

[3] *Sequential, as used here, implies that when the equations are properly ordered the solution of each may be based on initial condition and previous equation solutions without simultaniety.*

[4] *Kalman J. Cohen, Computer Models of the Shoe, Leather and Hide Sequence; Prentice-Hall, Englewood Cliffs, N. J., 1960.*

[5] *Guy H. Orcutt, et. al, Microanalysis of Socio-Economic Systems: A Simulation Study; Harpers, New York, 1961.*

[6] *Jay W. Forrester, Industrial Dynamics; Wiley, 1961.*

Chapter II

THE LAND USE PLAN DESIGN MODEL

THE LAND USE PLAN DESIGN PROCESS

To appreciate the need for, and requirements of, a Land Use Plan Design Model, it is necessary to examine closely the design process in general and the land use plan design process in particular. Analytical discussion of the design process is rare. Most of the literature on design is based on intuitive and artistic concepts or "styles" that have predominated in certain periods of history.

An exception to this general scarcity of literature is a recent work by Alexander,[7] in which he defines the design problem in terms of a "fit" between the problem statement and its solution.

> It is based on the idea that every design problem begins with an effort to achieve fitness between two entities: the form in question and its context. The form is the solution to the problem; the context defines the problem. In other words, when we speak of design, the real object of discussion is not the form alone, but the ensemble comprising the form and its context.

Achieving this fit between the form and its context is not a simple task since the many design requirements that make up the context of the design problem often interact in a complex manner. Attempts to satisfy one design requirement often leads to a violation of another. Faced with such complexity, a temptation exists for the designer to ignore the real design problem and substitute a traditional design that fails to solve the real design problem. Although such an approach may be acceptable in a political sense, the original problem remains unsolved.

Difficulties in the design process derive primarily from the inability of the human designer to simultaneously manipulate a large number of interacting design relationships. Mathematics, particularly in its newer forms, such as modern algebra, provides a powerful tool for the manipulation of these relationships for the more effective solution of design problems.

To be useful in design synthesis, mathematical formulations must comply with two conditions related to the foregoing definition of a selection problem:[8]

> It must be possible to generate a wide enough range of possible alternative solutions symbolically.
>
> It must be possible to express all criteria for solution in terms of the same symbolism.

While Alexander does not pursue the direct solution of selection problems using mathematical techniques, his definition provides a useful criteria for the systematic formulation of such problems.

Land use plan design, despite its admitted complexity, possesses certain inherent characteristics that meet Alexander's requirements of a "selection problem." The first requirement, involving the symbolic generation of a wide range of alternative solutions, is naturally achieved in land use plan design by reason of the common measure of all land use plans—the land itself. All land use plans for areas ranging from the smallest subdivision to multi-state

---

[7] *Christopher Alexander, Notes on the Synthesis of Form; Harvard University Press Cambridge, Mass., 1964.*

[8] *Ibid.*

regions may be symbolically expressed by three sets of variables: 1) the type of land use (quality variables), 2) the density of land use (quantity variables), and 3) the geographic location (location variables).

Typically, the land area concerned will be subdivided into a grid of "zones" of approximately equal area. The location variable is determined by the geographic coordinates of the zone in question. For each zone the types and densities of land uses may be expressed as a measure of the activities in that zonal area. The amount of detail provided will depend on the coarseness of the grid. For small areas a zone may be as small as individual residential lot parcels. In large regions they may be townships. The key point to be observed is that all land use plans may be expressed by these three classes of variables.

It should be stressed that the grid nature of the coordinate system does not confine the plan design to unimaginative, rectangular forms. On the contrary, the most complex and irregular plans may be expressed with the designated variables if an appropriate grid size is selected.

The second condition, relating to the symbolic relationship between alternative forms and design requirements, is also complied with in the land use plan design problem. All design requirements or "standards" restrict in some way the set of acceptable land use plans. From a design model point of view, these requirements subdivide into two primary classes: 1) requirements that restrict the minimum or maximum value of a density of land use or the relationship between land uses within a grid zone (intra-zonal standards), and 2) requirements that restrict the relationship between land uses between grid zones (inter-zonal standards).

In either class the design requirement can be expressed symbolically as an algebraic equality or more often an inequality relationship using the three classes of variables noted above. Again, compliance with this condition, like the first, is possible because land use planning is concerned with a single measurable resource—land. That the claims of symbolic design alternative generation and the requirements-alternatives comparison are authentic will become more apparent as the design model methodology is further explained.

From a practical standpoint, it is well at this point to provide a very specific succinct statement of the land use plan design problem. It is important to know the nature of both the design requirements and the design alternatives. To an experienced urban planner, the problem will certainly not be new, since it is the same basic problem he has been intuitively concerned with during his past design experiences. The problem, as it is stated, may seem somewhat excessively quantitative; and the emphasis on minimal costs may appear unnecessary, but fundamentally it is the same problem of urban form design that has challenged man since cities were found useful.

In brief, the problem of the urban land use design problem is:

1) Given design requirements expressed as:
   a) A set of design standards in terms of restrictions on land use relationships that may exist in the plan.

   b) A set of needs or demands for each type of land use based on forecasts of future urban activity.

2) Synthesize a land use plan design that satisfies both the land use demands and design standards considering the current state of both natural and man-made land characteristics at a minimal combination of public and private costs.

The conceptual basis for minimal costs, it must be emphasized, is not to provide a "cheap" plan but to avoid unnecessary expenditures of precious resources and at the same time comply with the design standards and land use demands in the plan design.

Intra-zonal design standards may take the form of limitations on density or restrictions of the types of land use that may coexist within a zone. An example of an intra-zonal design standard would be the provision of a regional shopping center within a certain travel time of every residential area. Land use demand requirements would restrict the set of acceptable plans to those that provide the aggregate total of each land use need over the entire design area. The current state of the land, be it developed or in a natural state, is a primary consideration in plan synthesis because of the relationship of the land to both the design standards and the costs associated with new or renewed development.

THE DESIGN MODEL

A Linear Programming Formulation

Two related mathematical techniques will now be discussed as possible frameworks for a Land Use Plan Design Model. The first technique, linear programming, has a record of successful accomplishment in other fields and has efficient, highly developed computational procedures. Dynamic programming, the second and newer technique, while not as productive in previous applications or standardized computational procedures, is less restricted in its assumptions and, potentially at least, is a more flexible framework for a Land Use Plan Design Model.

Both linear and dynamic programming are sometimes classified as subsidiary fields under the general title of mathematical programming. Such a general classification is desirable, inasmuch as both fields have as their objective the solution of problems involving the optimization (maximization or minimization) of some objective; for example, cost within the restrictions of certain constraints, such as design standards. The techniques involved differ considerably, however, with linear programming imposing rather severe restrictions on the nature of both the objective and constraints, while dynamic programming is almost unrestricted in its formulations of both the objective and constraint functions. Linear programming models, on the other hand, can usually be solved, once formulated, using standardized computational procedures, while dynamic programming usually provides at least a serious challenge and often insurmountable obstacles to an efficient computational solution. With either technique, the sheer size of many land use plan design problems brings with it what has been called the "curse of dimensionality" which militates against any straightforward "brute force" approach to solution.

The linear programming formulation of the Land Use Plan Design Model problem is straightforward. The objective function relates to the cost of developing land for a given land use.

$$c_t = c_1x_1+c_2x_2+\dots+c_nx_n$$

where the variables (x) may represent residential, industrial or other land uses in given areas and the constants (c) the costs of developing this land. Land use categories may be subdivided into subsidiary classes, such as single-family residential or multi-family residential, and the costs may be related to the topography and soil characteristics of the area.

With each subdivision, of course, the number of variables grows larger; and the computation time for a model solution is increased. In practice, a compromise must be made between the desire for detail and reasonable solution times. With the rapid development of computer technology, however, this problem will become of decreasing significance.

The equality and inequality constraints in the land use plan design linear programming formulation include the following:

1) The total demand requirement for each land use category.
(equality constraint)

$$d_1x_1+d_2x_2+\ldots+d_nx_n = E_k$$

where

E - regional land use demand requirement for each land use.

* d - service ratio coefficients which provide for supporting service land requirements, for example, streets which are necessary for primary land use development.

2) Maximal (minimal) limits on land uses within a zone.

$$x_1+x_2+\ldots+x_n \leq F_m$$ (usually density constraint)

where

F_m - Upper limit on land use n in zone m

3) Inter-zonal or intra-zonal land use relationship constraints.

$$x_n \leq Gx_m$$

where

G - Ratio of land use n allowed relative to land use m with land uses m and n in the same or different zones

The land use demand equality constraint follows a standardized format with one equation for each primary land use category. Since some land uses, such as single-family residential, are usually subdivided further according to lot sizes, the number of demand equations in a typical design model may exceed twenty relationships. It is important to emphasize that only primary land uses, such as residential, industrial, agricultural, and recreational land, need be directly determined. Service ratios incorporated in the d parameters can account for secondary land uses, such as local streets and parks.

The second and third categories of constraints reflect the design standards and may take a wide variety of forms. The maximal constraint will usually reflect a density standard, but it also may provide for the exclusion of an unsuited soil type area for a given type of land use. Land use relationship constraints will result from design standard restrictions on coexistent land uses within a zone or in adjacent zones. Accessibility standards for employment and shopping areas will also be reflected in this type of constraint.

The above constraint relationships reflect the types encountered so far in experimental plan design model runs in test areas. Other constraint forms may be needed when a complete regional plan design is attempted; but they may be easily included as long as they are linear, continuous constraints. Nonlinear, discontinuous constraints, such as the need for subdivision units of varying size, are not possible with linear programming and account for the primary disadvantages of the method. These problems and their solution by a second mathematical discipline, dynamic programming, will be briefly discussed in the latter part of this section.

For a region subdivided into about thirty zones, the size of a typical linear program for a land use plan design is about 60 equality and inequality constraints and 400 variables. Computer time on an IBM 1620 computer is about three hours to calculate the plan design. On larger systems, such as the IBM 7090, computer time is less than thirty minutes.

Although linear programming provides a reasonably satisfactory framework for a land use plan design model, it possesses certain inherent disadvantages that restrict its usefulness in design. The primary limitation is the need for continuous rather than discrete values for the land use variables. Land use design choices are by nature usually discrete rather than continuous. The basic element of residential land use development is the subdivision rather than the lot. Industrial land use units tend to be industrial parks rather than individual factory sites. While it is possible to round off the linear programming solution to satisfy these natural discrete levels, such a solution does not usually correspond with the associated discrete optimal combination.

A second limitation of linear programming is the need for both a linear objective function and linear constraints. The linear objective function is not a severe limitation because the inaccuracies introduced by a linear approximation of costs are usually less than the errors of cost estimation. In the few instances where known nonlinear cost functions occur, such as the plan capacities of areawide facilities for water supply or sewage treatment, the cost break may usually be satisfactorily approximated by a multivariable series of linear cost variables.

Nonlinear constraint relationships present a more serious problem. Certain design standards are inherently nonlinear, and a linear approximation sometimes provides an unsatisfactory substitute. When a design model is not able to satisfactorily provide for a design standard, it loses most of its usefulness.

An Alternate Dynamic Programming Formulation

Dynamic programming, another member of the mathematical programming family, has the potential for removing the two primary restrictions inherent in linear programming. Although dynamic programming may be used to solve the same land use plan design problem, it is based on a different class of algorithms, which are capable of handling discrete and nonlinear objective functions and constraint relationships.

Richard Bellman of the Rand Corporation was the originator of dynamic programming and has developed the theory and application of this multi-stage approach to decision-making to a high degree in the last decade. A large number of classes of dynamic programming processes have been formulated in areas such as production scheduling, rocket trajectories, and feedback control systems; ~~but the class of process of primary interest in design is the allocation process.~~[9]

~~The simple one-dimensional allocation process is formulated to obtain the maximum (minimum) of n variables:~~

$$R(x_1, x_2, \ldots, x_n) = g_1(x_1) + g_2(x_2) + \ldots + g_n(x_n)$$

over the Region constrained by the relations

$$x_1 + x_2 + \ldots + x_n = x$$

$$x_i \geq 0$$

~~where R is the return obtained from allocating a total resource, such as land, to different activities.~~ It is apparent that this problem is equivalent to a linear programming problem with a single equality constraint. The difference lies in the nature of the return functions, $R(x_n)$. Unlike linear programming, the values of x_n may be discrete; and the corresponding value of the return $R(x_n)$ may be as nonlinear as required. Because the basic algorithm is numeric rather than analytic, there are almost no restrictions on the nature of the return function.

A "brute force" numerical approach to the above problem might consider the direct enumeration of all cases. There is no question that the optimal return may be determined by calculating the returns for all possible combinations of the variables involved. A brief example, however, will illustrate the impracticality of this approach. If each variable x_n al-

---

[9] *Richard E. Bellman and Stuart E. Dreyfus, Applied Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1962.*

lowed for ten different levels, and there were ten activities, the number of cases are 10^{10} or ten billion. On the fastest present day computers, allowing only a millisecond for each enumeration, computation time is of the order of ten years. Even with the rapid advances in computer technology, it is evident that the direct enumerative approach is not practical, particularly since the above problem is simple compared to a realistic plan design problem.

The basic algorithm of resource allocation in dynamic programming significantly reduces the computation effort using the following procedure:

$$f_n(x) = \underset{0 \leq x_n \leq x}{\text{Max}} \left[g_n(x_n) + f_{n-1}(x - x_n) \right]$$

$$\text{For } n = 2, 3, \ldots, N \qquad x \leq 0$$

and

$$f_1(x) = \text{Max}[g_1(x)]$$

This algorithm illustrates the Principle of Optimality which states:[10] "An optimal policy has the property that whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

This principle, while easy to state, is difficult to grasp at first; and an example may be helpful to illustrate its application. Using the land use plan example and assigning land to multiple uses n within a given total land area x so as to minimize the development costs, dynamic programming may be explained as a multi-stage decision process.

First, determine the minimal cost for each level of land assigned to the first land use. This first solution, $f_1(x)$, is trivial since it will simply be the minimum of the return function, $g_1(x_1)$.

The second stage decision, $f_2(x)$, determines the minimal cost for the first and second land uses in combination.

$$f_2(x) = \underset{0 \leq x_2 \leq x}{\text{Min}} \left[g_2(x_2) + f_1(x - x_2) \right]$$

The third and succeeding stage decisions are then computed using the preceding decision function, $f_{n-1}(x-x_n)$. The third decision function, $f_3(x)$, minimizes the first three land uses; and succeeding functions minimize the combination of the land uses up to the n stage. When the final land use has been included in the final stage, a minimal combination of all land uses for each level of land in the area will have been determined.

With the series of decision functions, $f_1, f_2, \ldots, f_n$, available, it is now possible to solve a reverse solution to determine the actual amount of each land use to be assigned for an optimal plan. Beginning with f_n, x_n may be determined by entering the function with the total land area, x.

The next stage of the reverse solution, f_{n-1}, is then entered with the land available after x_n was assigned, $x-x_n$.

The reverse solution proceeds through the remaining stages, $f_{n-2}, \ldots, f_1$, until all of the land in the area is assigned and an optimal plan is determined.

[10] *Ibid.*

Although the above example illustrates the above basic principle involved in dynamic programming, it is an oversimplified view of a land use plan design model in that only a single constraint, the total land area, is involved. A typical land use plan design problem involves the assignment of land to multiple land uses in multiple zonal areas under the restrictions of multiple design standard constraints.

The typical land use design problem may be restated:

Minimize:

$$R(x_1,x_2,\dots,x_n) = g_1(x_1)+g_2(x_2)+\dots+g_n(x_n)$$

$$x_i \geq 0$$

Subject to:

$$\sum_{i=1}^{k} \left(\sum_{j=1}^{n} a_{ij}x_j\right) \leq x_i$$

where the final restrictions represent all of the constraints of the problem, such as the total land use demands and land use relationships previously formulated in the linear programming model.

The relatively simple and straightforward dynamic programming approach just described for a single-constraint solution is not computationally efficient for the multiple-constraint problem. The use of functional equation relationships does not sufficiently reduce the dimensionality of the problem. The number of computations required tends to expand by the power of the number of constraints and becomes inefficient for more than two constraints. While such an approach is to be preferred over direct enumeration, it still lies beyond the range of economic feasibility with present day computers.

Bellman provides a possible way out of this dilemma through the use of a mathematical device originally developed in calculus but actually of wider significance, the Lagrange multiplier. A complete discussion of the Lagrange multiplier is set forth in the Bellman and Dreyfus text[11].

The practical significance of the use of the Lagrange multiplier in the Land Use Plan Design model is the transformation of the original objective function to be minimized.

$$R(x_1,x_2,\dots,x_n) = \sum_{j=1}^{n} g_j(x_j) - \sum_{i=1}^{k} \lambda_i\left(\sum_{j=1}^{n} a_{ij}x_j\right)$$

where λi are the Lagrange multipliers. In the revised formulation, a series of single-constraint dynamic programming solutions are computed with varying values of the Lagrange multipliers until all of the constraints are satisfied.

The basic feasibility and practicality of the dual constraint dynamic programming problem using a Lagrange multiplier has been demonstrated, and experience with various extensions of the Cord[12] model, which uses the Lagrange multiplier technique, has been gained within the Region in industrial management problem applications.

---

[11] *Ibid.*

[12] *J. Cord, "A Method for Allocating Funds to Investment Funds When Returns Are Subject to Uncertainty"; Management Science, January 1964.*

Experimentation with the dynamic programming version of the Land Use Plan Design Model is still in the early stages, but initial results are encouraging. One of the important findings in early experience with the model is that many design standards may be formulated implicitly in the return function for the land use activity or by the omission of designated activities in certain areas. Such a formulation greatly reduces the number of constraints that must be determined by means of the Lagrange multiplier.

Initial model tests have also resulted in a more rapidly converging version that uses two types of Lagrange multipliers:

$$R(x_1,x_2,\ldots,x_n) = \sum_{j=1}^{n} \lambda_i' g_j(x_j) - \sum_{i=1}^{k} \lambda_i \left(\sum_{j=1}^{n} a_{ij} x_j \right)$$

where the additional multiplier set, λ_i , is multiplicative rather than additive. This new multiplier has been used for areawide constraints, such as the total demand for land use in the Region. The additive multiplier, λ_i , has been used for the intra-zonal constraints. Although experimental experience is still too limited to draw conclusions, this approach seems to expedite convergence of the solution.

Although linear programming offers a satisfactory vehicle for a limited Land Use Plan Design Model that may later be extended through the use of newly developing integer and nonlinear programming techniques, the dynamic programming model offers such flexibility and power that its ultimate success could have an extremely beneficial effect on urban plan design.

LAND USE ACTIVITY CLASSIFICATION

Two classes of land use were used in the Land Use Plan Design Model. Primary land uses are directly assigned to zonal areas by the model in its constrained-optimization procedure. Secondary land uses are indirectly assigned since they are determined as a result of the primary allocation. The following primary and secondary land use classifications were used for initial experimentation with the Land Use Plan Design Model.

1) Primary Land Uses.

   a) Residential
   b) Regional Retail Trade and Services.
   c) Manufacturing and Wholesale Trade.
   d) Regional Recreation.
   e) Agriculture.

2) Secondary Land Uses.

   a) Local Retail Trade and Services.
   b) Transportation, Communication, and Utilities.
   c) Institutional and Governmental Services.
   d) Local Recreation.

The availability of cost data and detailed design standards in this experimental phase of the design model program limited consideration to the following land use categories:

3) Primary Land Uses (experimental tests).

   a) Residential.
   b) Regional Retail Trade and Services.

4) Secondary Land Uses (experimental tests).

a) Local Retail Trade and Services.
b) Local and Collector Street Right-of-Way.
c) Local Recreation.
d) Local Institutional and Governmental (schools).

The above experimental test classification, although limited, provided sufficient latitude to "exercise" the basic operation of the model since the location of regional shopping centers introduced inter-zone as well as intra-zone design standards. An accessibility standard was imposed that required all residential areas to be within a specified travel time of a regional shopping center.

It should be emphasized that the above classification is not to be considered rigid or final. Some land use categories classified as secondary, such as institutional and governmental services, could be reclassified as primary. Other secondary land uses like transportation, communications, and utilities will ultimately, in the operational version of the model, be a combined primary-secondary category since the cost of these facilities should be considered concurrently with the assignment of primary land use to preserve the cost optimality of the plan. Such an extension of the model required cost and design standard data not available in the present limited development program. Later extensions of the model are planned to provide for this joint optimization.

To provide for a more detailed land use plan design, residential land use was further subdivided into five residential lot sizes for single-family homes. For each of these lot sizes, detailed development costs and secondary land use design standards were provided. A summary of these costs and standards is tabulated in Tables I and II. Two types of shopping centers, major and minor, were included under the regional retail trade and services land use.

MODEL INPUT DATA REQUIREMENTS

Previous sections of this report have defined the role of the Land Use Plan Design Model in the land use planning process, introduced the conceptual design framework for the model, described two computational techniques used in model implementation, and established the land use classification. In this and the following section, some initial experience with applications of the model will be detailed in order to provide an indication of the input data requirements and computational characteristics of the model. It should again be emphasized that the application of this design model is still in the embryonic stages, and no definite design recommendations should be implied from any of the experimental test results.

Four primary sets of input data are required for model operation:

1) The costs of unimproved land and land development for each primary land use activity for each type of soil.

2) The aggregate demand for each primary land use activity.

3) Design standards which reflect the plan objectives and restrict the set of acceptable plans by limiting inter-zonal and intra-zonal land use relationships.

4) A current land inventory which will include both existing land use activities by area and soil characteristics.

Land development cost data may be obtained either by engineering estimates made by personnel familiar with land development or by statistical analysis of recent land development in the area. The former approach has been used in the initial tests of the model in the Waukesha City pilot area. Collection of land development cost data is always expensive and in many cases difficult or even impossible to obtain. Land developers are usually extremely reluctant to reveal their costs, and the cost data obtained is of uneven quality since many developers do not maintain complete records. For all of these reasons, engineering cost estimates are usually preferable if competent professional experience is available.

In the Waukesha area, separate land development cost estimates were made for the five sizes of residential lots with their associated service land uses, such as streets, neighborhood shopping, schools, and parks. Additional cost estimates were made for industrial, regional shopping, and regional park land uses. These cost estimates were not just gross estimates but detailed analyses of the costs of each improvement related to both the land use and the type of soil involved. All cost estimates were subdivided into their component parts, each with its own individual cost estimate. Separate cost estimates were prepared for each of three

Table I
RESIDENTIAL PRIMARY AND SECONDARY LAND USE CATEGORIES, DESIGN STANDARD SUMMARY
(residential module)[a]

Neighborhood Characteristics	Lot Size (square feet)				
	Under 9,000	9,000 to 11,000	12,000 to 19,000	20,000 to One Acre	Over One Acre
Average Size (feet)					
Lot Frontage	65	80	100	150	300
Lot Depth	120	130	150	200	300
Block Length	900	900	1,200	1,200	1,500
Residences					
Single-Family					
Percent of Total Area	62.0	64.0	67.2	76.5	88.3
Number of Lots	2,240	1,710	1,180	710	272
Multi-Family					
Percent of Total Area	6.0	5.0	7.0	1.0	---
Number of Lots	203	144	129	10	---
Estimated Population	8,730	6,590	4,750	2,480	900
Percent of Total Area for:					
Parks	3.0	2.5	2.0	1.5	1.0
Schools	1.5	1.5	1.0	1.0	0.5
Retail Services	1.0	1.0	0.8	0.5	0.2
Streets	24.0	24.0	20.0	18.0	9.0
Other	2.5	2.0	2.0	1.5	1.0

[a] *Module area varies by density class as follows: low density; 2 miles square; medium density, 1 mile square; high density, 1/2 mile square.*

Lot Type

1 640 Acres (1 square mile).
2 Single Family 3.3 Persons/Lot.
3 2 Family/Lot 6.6 Persons/Lot.
4 4 Family/Lot 13.2 Persons/Lot.
5 15% (0.5-family) Elementary School Age.

Table 2

ESTIMATED LAND IMPROVEMENT COSTS BY LOT TYPE

IMPROVEMENT	COSTS ON		
	Very Good and Good Soils	Fair Soils	Poor Soils
Street and Utility Improvement			
Lot Size			
Under 9,000 sq. ft.	$3,560.00/lot	$3,945.00/lot	$5,050.00/lot
9,000 to 11,999 sq. ft.	4,345.00/lot	4,805.00/lot	6,149.00/lot
12,000 to 19,999 sq. ft.	5,443.00/lot	6,003.00/lot	7,682.00/lot
20,000 sq. ft. to 1 acre	3,640.00/lot	4,827.00/lot	6,477.00/lot
Over 1 acre	5,608.00/lot	5,894.00/lot	7,298.00/lot
Neighborhood School Construction			
Lot Size			
Under 9,000 sq. ft.	$ 769.00/lot	$ 772.00/lot	$ 822.00/lot
9,000 to 11,999 sq. ft.	658.00/lot	661.00/lot	707.00/lot
12,000 to 19,999 sq. ft.	442.00/lot	449.00/lot	478.00/lot
20,000 sq. ft. to 1 acre	687.00/lot	696.00/lot	745.00/lot
Over 1 acre	659.00/lot	661.00/lot	711.00/lot
Neighborhood Park Improvement			
Lot Size			
Under 9,000 sq. ft.	$ 76.00/lot	$ 82.00/lot	$ 106.00/lot
9,000 to 11,999 sq. ft.	83.00/lot	89.00/lot	118.00/lot
12,000 to 19,999 sq. ft.	97.00/lot	104.00/lot	135.00/lot
20,000 sq. ft. to 1 acre	87.00/lot	106.00/lot	141.00/lot
Over 1 acre	143.00/lot	146.00/lot	185.00/lot
Neighborhood Commercial Center Construction			
Lot Size			
Under 9,000 sq. ft.	$ 572.00/lot	$ 631.00/lot	$ 707.00/lot
9,000 to 11,999 sq. ft.	377.00/lot	416.00/lot	466.00/lot
12,000 to 19,999 sq. ft.	534.00/lot	589.00/lot	660.00/lot
20,000 sq. ft. to 1 acre	460.00/lot	515.00/lot	576.00/lot
Over 1 acre	605.00/lot	666.00/lot	737.00/lot
Total (Combined Improvements)			
Lot Size			
Under 9,000 sq. ft.	$4,977.00/lot	$5,430.00/lot	$6,685.00/lot
9,000 to 11,999 sq. ft.	5,463.00/lot	5,971.00/lot	7,440.00/lot
12,000 to 19,999 sq. ft.	6,516.00/lot	7,145.00/lot	8,955.00/lot
20,000 sq. ft. to 1 acre	4,874.00/lot	6,144.00/lot	7,939.00/lot
Over 1 acre	7,015.00/lot	7,367.00/lot	8,931.00/lot

classes of soil. Soil data was obtained from a detailed operational soil survey of the Region conducted by the United States Department of Agriculture, Soil Conservation Service in co-operation with the Southeastern Wisconsin Regional Planning Commission.

Unimproved land costs presented a special problem since they could not be obtained from engineering estimates. Assessed and equalized land value data was obtained from the state equalized assessment roles of each of the communities and was adjusted based on the prices realized in recent land transactions in the area.

Initial experimental tests of the model used historical aggregate land use demands for the 1950-1962 time period to provide comparisons between actual and "optimal" land development in the area. Typically, however, a design application will require forecasts of future land use demands, which are obtained by applying design standards to forecasts of population and exployment in the region of interest, as described in the first section of this report.

The various forms of design standards usually provided were described in the previous section. In experimental model tests, design standards have been limited to the exclusion of certain areas from development, such as flood plains and wet soils, along with the provision of service ratios for the amounts of secondary land, for example, streets, required to support the primary land uses. Design standards for the regional land use plan are still in preparation and will be used in later model tests as soon as they become available.

An inventory of both current land use activities and soil characteristics is critical for model application. In current tests developed areas were eliminated from consideration for future land development. It is possible, however, to reconsider redevelopment in the form of urban renewal as a set of alternatives in the design. For this approach redevelopment costs would be required. Through the use of the detailed soil inventory, it was possible to assign a development cost to each subarea in the test area.

All of the data input eventually manifests itself as a parameter in the cost vector, constraint vector, or requirements matrix of the linear programming model. Although the input format differs, the same class of inputs apply to the dynamic programming model.

MODEL APPLICATION

The results of applying the design model to the City of Waukesha and its environs, using the historical (1950-1962) aggregate land use demands as the demand requirements and the standards shown in Table 1, as the design standards, are illustrated in Figure 2. It is interesting to contrast the model design with the actual land development in the Waukesha area during this period. As might be expected, the actual development was considerably more scattered and fragmented than its model design. Of special interest is the design standard requiring all of the three smaller lot sizes to be provided with sanitary sewer. Although this design standard was generally complied with in the actual land development, it is still noteworthy because of the singular effect of this one standard on the land development pattern.

It would be fallacious to interpret these experimental model results as a critique of the plan for the City of Waukesha. The model is still too experimental to serve as the basis for actual design. Indeed, the circumstances surrounding actual land development are appreciated since the Land Use Simulation Model, to be described in the latter half of this report, was tested in this same area and time period. These simulation tests provided some insight into some of the practical constraints that must be faced in implementing an ideal land use plan design.

Although the land use classification was limited, the costs preliminary, and the design standards rudimentary, the experimental tests of the Land Use Plan Design Model were important in establishing the basic feasibility of the design model. All of the important types of costs and design standards were tested except those associated with arterial highway and central utility facilities. Although the potential of the model became obvious to all participants in its development, its problems and needs for further development were also noteworthy.

MODEL EXTENSIONS AND NEED FOR FURTHER DEVELOPMENT

At this early stage of design model development, it is not possible to provide an exhaustive enumeration of model development needs. Such needs will become more obvious as the model approaches operational status. Early experimentation and critical discussion have, however, revealed the more obvious needs and problems.

Most important, perhaps, is the requirement to jointly consider the costs of transportation and central utility facility extensions concurrently with the allocation of the primary land uses. These facility costs are often so critical and major in the economics of land development that to ignore them is to severely distort the nature of the model design. A promising approach providing for joint optimization of land development and major facility costs, using the Moore algorithm of traffic assignment fame, is now being evaluated. If such joint optimization proves practical, the Land Use Plan Design Model may be extended to become a land use-transportation plan design model.

Another area of difficulty relates to the manufacturing and wholesale trade land use categories. Although these land uses were not assigned in the experimental model tests, they would have to be included in any operational model. The requirements of different industries are diverse, and it is not yet clear whether any meaningful aggregation of industrial land uses is possible.

Related to both of the problems and other extensions of the model is the general need for more substantive research in all areas of the model. At this point in time, it appears that the mathematical and computational capabilities of the model are far ahead of the substantive knowledge supporting it.

Substantive needs of the design model, although primary at this time, should not obscure the very real needs for mathematical and computer program development. It is apparent, even at this early stage, that dynamic programming will provide the ultimate model algorithm. Much experience needs to be acquired to make such promise a reality.

The nature of the remaining model needs and research requirements would indicate the need for a research team embracing an experienced urban planner and civil engineer, as well as a systems engineer. Future progress will depend on the ~~optional~~ optimal mixture of analytic and substantive knowledge.

Chapter III

THE LAND USE SIMULATION MODEL

MODEL OBJECTIVE

The objective of the Land Use Simulation Model is to provide a means of testing regional land use plans for feasibility of implementation. The emphasis is not on forecasting but on plan implementation. The model is intended to test the effectiveness of certain controlled variables in achieving a given target plan in the presence of many uncontrolled variables. Controlled variables will represent the implementation tools of land use planning: public land use controls, public facilities construction, and public land acquisition. Uncontrolled variables will include the behavior of households, private land developers and builders, and exogenous inputs, such as population growth and employment.

Although the primary use of the Land Use Simulation Model will not be in forecasting, one of the applications of this model in southeastern Wisconsin will be a simulation of current trends in the regional land use pattern given the existing public works programs and land use controls in the Region. In one sense, such a simulation is a forecast since none of the public control variables would be affected by the regional plan. The purpose of this simulation is to present for public consideration the questionable desirability of the emerging land use pattern without a comprehensive regional plan.

Most of the land use simulations, however, will be concerned with the experimental design of policy to implement a target land use plan. The end product will be a set of public works programs and land use regulations needed to achieve the regional land use plan.

MODEL ORGANIZATION

The Land Use Simulation Model is a dynamic behavioral feedback simulation model and is classified into five primary sectors:

1) Residential.
2) Industrial.
3) Services.
4) Special.
5) Agricultural.

In the residential sector, the decision-making behavior of "household-type" units are simulated in conjunction with the related decisions of land developers and builders. Variables influenced by the land use planner, as later reflected in governmental policies, are programmed to achieve the desired land use pattern. These controls tend to constrain or modify the behavior of households, land developers, and builders.

The industrial sector in current model tests is being treated exogenously, with industrial employment in each zone being programmed in light of the land use plan. A second experimental endogenous version of the sector is now being tested for later incorporation in the model. In this latter approach, "firm types" determined from an industrial classification select new industrial sites based on their particular requirements and the costs of land and taxes. Although the endogenous approach to industrial location simulation has a certain appeal in that it provides a behavorial explanation of industrial location decisions, the exogenous approach may be more in keeping with the planning approach described earlier in this report. If the sites of industrial employment are a powerful influence on residential and service-related land development, then implementation of the target plan will probably require a governmental influence on these decisions. If such influence can take the form of providing land with the characteristics needed by the various industrial groups at prices they are willing to pay, then the exogenous and endogenous versions of industrial land development should be similar.

The service sector of the model embraces all land uses, the location of which are primarily dependent on accessibility to residential and industrial land. Such land uses include not only local retail and service establishments but also schools, local streets, and neighborhood parks. A dual interdependency exists for some of the land uses in this category, such as retail trade and schools since their location is dependent on residential and industrial land use, but they also influence this same residential and industrial land use pattern in a feedback fashion.

The special sector includes all nonindustrial exogenous inputs to the model, most of which are the result of governmental decisions. These include the major freeway and arterial network, regional park and open space areas, and rail-utility rights-of-way and terminals.

Agricultural land use is treated in a residual manner in the model with such land being transferred to other land uses during the simulation period. Such a representation does not imply an endorsement of the gradual disappearance of agricultural land in the Region. In fact, such representation is intended to emphasize the need to consider the relative economic, resource conservation, and aesthetic worth of such land in the land use plan design and thus provide the need for the formulation of policies to prevent this conversion of agricultural land should it prove undesirable.

MODEL CHARACTERISTICS

It is convenient at this point to review some of the characteristics of the Land Use Simulation Model, particularly those that differ from other land use models being developed under the auspices of other agencies. The differences enumerated below should not be interpreted as a criticism of other model development in this field. The current experimental state of land use model development does not permit anyone to assert the absolute validity of a given conceptual approach. Then, too, planning objectives differ; and the Land Use Simulation Model under discussion may not be ideal or even useful in other planning programs. In the current embryonic state of land use models, alternative approaches, even if ultimately unsuccessful, should add to the store of research knowledge in the field.

The dynamic nature of the model has been explained previously and will not be belabored again, except to point out that many land use models are static in nature having been formulated to determine a land use pattern at a single point in time. Such a static approach, it is admitted, has usually resulted from data deficiencies rather than any basic disagreement about the desirability of a dynamic model.

A second important feature of the model is its degree of disaggregation. A more detailed model is consistent with a behavioral decision-making approach to model formulation. Since households differ considerably in their income, education, age, and other characteristics, the use of an aggregate household in the model is subject to question. For this reason, households have been classified into types with common characteristics, with the hope of obtaining stability in the model parameters. Further disaggregation has been accomplished by the subdivision of household relocational behavior into a number of subdecisions. Although disaggregation has its penalties in terms of additional data requirements, additional model segmentation was felt necessary to be consistent with the formulation of behavioral decision rules.

A sampling approach to parameter estimation was used in the Land Use Simulation Model. To implement this sampling approach, new data sources were required, including special household history data collected in the home interview part of the regional travel surveys. The use of this new household data will be described in a later section concerned with parameter estimation in the Land Use Simulation Model. Another important new data source, the detailed operational soil survey, also plays a critical role in the site selection decision of the land developer in the residential sector of the model.

Finally, the all encompassing characteristic of the model lies in its emphasis on the control rather than the forecasting function. Such emphasis is consistent with the generally accepted primary use of simulation models as vehicles for policy formulation. This "if-then" usage of a simulation model requires less information concerning the uncertain future values of exogenous variables than an equivalent forecasting usage. For this reason, conclusions may be drawn from model results with a higher degree of confidence.

RESIDENTIAL SECTOR RELATIONSHIPS

Recursive Programming

Primary emphasis in this section will be placed on an explanation of the residential sector of the model. This sector is fundamental to the operation of the model with the industrial, service, special, and agricultural sectors in auxiliary roles. The operation of the residential sector revolves about three primary decisions affecting the development of residential land: 1) the decision of the land developer to subdivide land for residential use, 2) the decision of the building contractor to build a dwelling unit or group of dwelling units, and 3) the decision of the household to rent or purchase a dwelling unit.

The above set of decisions, constrained by zoning, subdivision regulations, and other restrictions imposed by local governments, combine to determine the residential land use of the Region. The time sequence of the decisions is not necessarily in the order listed above. The household may dictate both the site development and house construction. It is also recognized that the household is the ultimate cause of the process since the sequence cannot continue if it refuses to buy or rent.

Every positivistic land use model is based upon an explicit or implicit theory of the land development process. Such a theory must be quite explicit in a behavioral decision-making approach to land use modeling since the decisive relationships in the model are a direct expression of the theory of the model.

~~The underlying theory of the Land Use Simulation Model is based on a land market modified by governmental policies in which households determine the quantity and composition but not the specific location of land and housing demand. The supply and specific location of housing sites are determined by the land developer according to his economic self-interest as limited by the information available to him.~~

Land and housing demand is dependent on the housing needs of a number of household types locating within the Region. The quantity of this demand is dependent on three sources of locators: 1) new household formations, 2) in-migrating households, and 3) households relocating internally within the Region.

The composition of this demand depends upon the needs of household types as determined from statistical analysis of historical household-housing combinations indicated by the household history survey conducted by the Southeastern Wisconsin Regional Planning Commission. Details of land-housing demand generation will be discussed subsequently after an introduction to the decision-making of land development.

Land development decision-making is formulated in the framework of a new approach to decision simulation, known as Recursive Linear Programming,[13] developed by Professor Richard H. Day, of the University of Wisconsin. This technique provides for "... optimized decision-making over a limited time horizon on the basis of knowledge gained from

1. what if no significant differences in land costs?
2. Don't hh demand housing by location?

[13] *Richard H. Day, Recursive Programming and Production Response; North-Holland Publishing Co., Amsterdam, 1963.*

past experience."[14] The analytical nature of the technique is best described in the words of its originator.

> Recursive Linear Programming is a sequence of linear programming problems in which the objective function, constraint matrix, and/or the right hand side parameters depend upon the primal and/or dual solution variables of the preceding linear programming problems in the sequence.

In the Land Use Simulation Model, the recursive programming relations take the following form:

$$\begin{cases} f(x) = \min cx(t) \\ A'x(t) = \hat{b}'(t) \\ A''x(t) \leq b''(t) \\ A'''x(t) \leq (1+B)x(t-1) \end{cases}$$

← cost of development
← demand for lot type

where, ~~$cx(t)$ represents the cost function minimized by the land developer while satisfying the demand for lots based on his estimate of the demand for each lot type, $\hat{b}'(t)$, and complying with the restrictions represented by the recursive constraint $A''x(t) \leq b''(t)$.~~

~~Expected demand for each lot type, $\hat{b}'(t)$, is some weighted function of the actual demand experienced in previous periods such as:~~

$$\hat{b}'(t) = \lambda b'(t-1)+\lambda^2 b'(t-2)+\ldots+\lambda^n b'(t-n)$$

$$\text{where} \quad \lambda<1$$

The actual demand vector during each period, of course, is determined by the household portion of the model. The above recursive linear programming relationships describe in essence the entire residential sector of the Land Use Simulation Model. The complete model relationships only elaborate the nature of:

1) Household demand generation of the $b'(t)$ vector.

2) Land developer estimation of demand $b'(t)$.

3) The constraint matrix which will include factors such as accessibility and zonal land capacities and can include behavioral adjustments of the land developer.

4) The cost parameters used in the model.

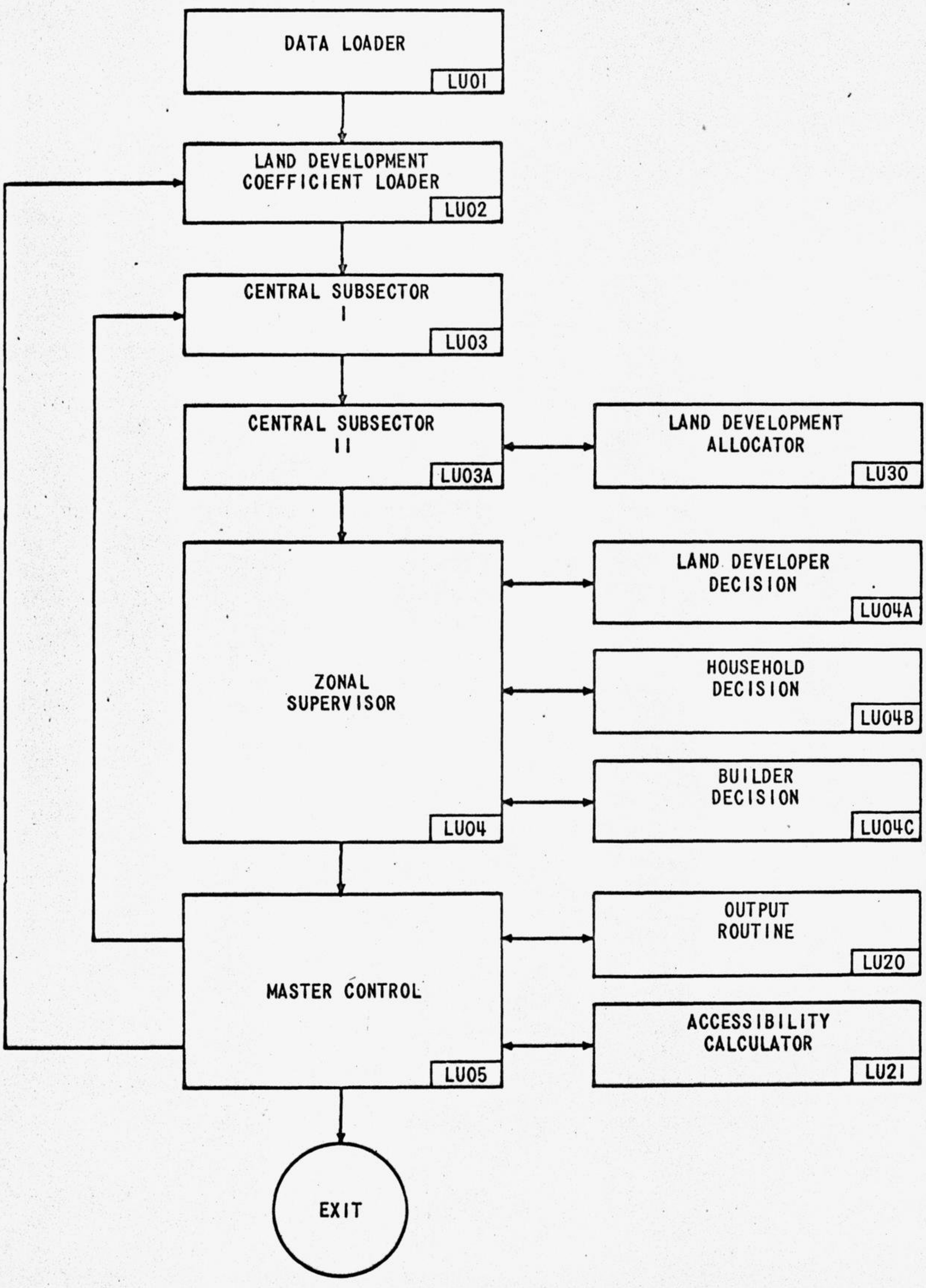
A subsidiary set of relationships for the housing contractor are also included in the model. These relationships resemble those of the land developer above. The prevalence of custom building with its associated demand orientation and the relative lack of speculative building in southeastern Wisconsin, however, make these builder relationships of lesser significance in the model.

Household Demand

The residential sector submodel, as illustrated in Figure 3, is subdivided into two subsectors: the central subsector and the zonal subsector. The central subsector is the central

[14] *Richard H. Day, Linear Programming and Related Computations; Department of Agriculture, Washington, D. C., 1964.*

Figure 3
LAND USE SIMULATION MODEL SYSTEM DIAGRAM



clearinghouse for all households relocating within the Region. New households, in-migrating households, out-migrating households, and households transferring from one zone within the Region to another one are temporarily located within this subsector until they are relocated at their new location. New households originate from an outside input to the model determined from population forecasts developed in the regional economic simulation model. In-migrating households are also an outside input from the economic model. All three of these relocating household sources are allocated to either one of the other zones or to out-migration.

The other subsector of the residential subsector is the zonal subsector. A zone is an area of land, varying in size from a U. S. Public Land Survey quarter section (160 acres) to a township (36 square miles), that will provide the basic areal unit in the model. At the present time, it is intended that the transportation origin and destination study analysis zones will be used wherever practicable. Some modification will undoubtedly be required in certain areas, but all modifications will be formed in terms of the basic areal data unit, the U. S. Public Land Survey quarter section.

To understand the zonal subsector, it is necessary to explain the concept of a household type. The validity of any system simulation model depends on the stability of the variable relationships and parameters used in the model. These relationships and parameters must be stable over time and over a range of outside inputs to the model. Experience tells us that households vary considerably in their behavior, depending on the characteristics of the household. The concept of an "average household" is elusive since any behavior peculiar to such a household is the result of combining a wide variety of units of differing behavioral patterns. The relationships and parameters used in a model for such a household will probably not be stable since the composition of an "average household" is continually changing because of changes in the number of different types of households.

To provide the necessary decomposition of households into types, all relationships in the model are subclassified by household type. A separate set of relationships and parameters are determined for each household type in the model. ~~In the zonal subsector equations, the number of households in each household type is recorded. This number is modified as a result of: 1) incoming households from the central subsector, 2) departing households relocating in a new zone or out-migrating, 3) aging households being transferred to the next older household type or being received from the next youngest household type, and 4) other household transitions resulting from changes in income or education.~~

Incoming households to each zone are determined by the number of households designated for relocation in the central subsector, together with the nature of developed land and housing available in the zone. The quantity and quality of available land and housing as limited by the accessibility of the zone to employment, shopping, and population influence the overall zonal growth for each household type.

Household locational decisions may be subdivided into three subdecisions: 1) the decision to move, 2) the selection of a land-housing unit type package, and 3) the selection of a particular site location.

This subdivision of the "where to move" decision into two subordinate decisions involving housing unit type selection and subsequent locational preference seems a natural one and is useful in providing statistical verification of the accuracy of the model representation of this decision process.

In model operation a relocating household of a certain type will leave its originating zone and be transferred to the central subsector, where it will be matched with an available housing unit type consistent with the locational preferences of the moving household.

To provide a classification of household types useful in simulating the above decisions, a special kind of statistical analysis of regional household data is required. The primary basis selected for household type classification was the housing unit type preference. Essentially, this involves matching household characteristics, such as age (of head), income, race, family size, and education (of head), with housing unit characteristics, such as structure type, market value of house and land, and owner or renter status. Other characteristics, such as water and sewer service and lot size, may be added to the housing unit type pattern as necessary.

A technique known as taxonomic analysis was used to determine the household and housing unit type classification. The criterion for classification was a good match between the household and its housing type preferences. In a good classification, each household type would be distributed among a limited number of housing unit types.

In the taxonomic approach, the quality of the match would be measured by the aggregate similarity ratio, which is defined as the ratio of the number of a household type included in a group of housing unit types. In a typical case, a household type might match four housing unit types with an aggregate similarity ratio of 0.96. This means that 96 percent of this household type was included in these housing unit types. The other samples might be scattered in a number of other housing unit types too numerous to include in the classification. The aggregate similarity ratio could always be improved, of course, by adding more housing unit types; but an unwarranted number of housing unit types of low-density representation do little to improve the usefulness of the match. The household and housing types used in the Land Use Simulation Model are described in Appendix II.

The turnover rate (decisions to move per month) can then be determined for each of the household types. Ideally, the variance of the turnover rate within each household type will be small.

After the relocating household is matched to a housing unit type, the site selection subdecision must be made. This subdecision is based on an accessibility constraint for each household type. The matched households will be distributed to zones with appropriate housing unit type vacancies that comply with the accessibility constraint for the particular household type. The concept here is that, once the household has decided to move and has selected its housing unit type, site location will be based on geographical considerations of accessibility to work, shopping, and other population groups. This use of accessibility differs significantly from the ordinary use of accessibility in regression-type models where the total change in population in a zone is related to accessibility alone. The approach here differs in the following respects:

1) Accessibility is a limited decision factor since it comes into play only as a constraint after the decision to move has been made and a housing unit type has been selected.

2) The value of the accessibility constraint will be varied to account for the differences in the importance of this factor for each household type.

The logic of the approach is clear in that the only consideration remaining after the decision to move and housing unit type selections have been made is geographic accessibility. Any remaining unexplained variance in the pattern should be the result of an incomplete formulation of the housing unit typology or a random element in the decision not capable of further explanation using this approach. In the actual case, some unexplained variance will remain. If this variance is less than 10%, it will not seriously jeopardize the usefulness of the model since this variance may be explicitly formulated in the simulation model.

It is possible that the site selection subdecision could be treated in a more microscopic fashion should it prove desirable. The information in the household history form used to collect the basic data will permit a more detailed analysis of employment accessibility by subareas and industries since that accessibility could be treated as another household type characteristic.

Model Implementation of the Household's Turnover Subdecision: The turnover subdecision is very simply implemented in the model. A turnover rate (TO XX) is associated with each household type (HXX), and the households of each type departing from each zone in each simulation period are the product of this turnover rate and the number of households of that type in the zone.

```
HD51(I) = [TO(51)][H51(I)]

Where

HD51(I) - Households of type 51 departing from zone I

TO(51)  - Turnover rate, household type 51

H51(I)  - Households, type 51, zone I
```

The turnover rate may be alternately expressed as a normal distribution with an average turnover rate (TOA51) and a standard deviation (TOD51).

Model Implementation of the Household's Housing Unit Type Selection Subdecision: The model implementation of the housing unit type selection subdecision is not as direct as that just described for the decision to move. Each zone will contain a limited number of housing unit types and for that reason will provide housing for only a limited number of household types. Households are related to the housing preference in the model through the use of a household housing matrix.

The existence of housing unit types in a zone is dependent on the decisions of the land developer and builder which, in turn, are dependent on zoning, subdivision regulations, the topography, soil, and other physical and social characteristics of the area. These decisions will be discussed in a later section.

Model Implementation of the Household's Site Selection Subdecision: Site selection in the model is based on the constraint of geographic accessibility. For each zone an accessibility factor for employment, population, and shopping is calculated, based on the travel time from the particular zone to all other zones and the relative attractiveness of the other zones in terms of their total employment (employment accessibility), population (social accessibility), and retail employees (commercial accessibility).

Travel times between zones are inserted as an outside input to the model and may represent any existing, historical, or proposed transportation network. These times may be changed during program operation to account for planned (or historical) changes in the network. Attractiveness factors will be based on the current status of the employment, population, and retail employment variables in the model.

The Land Developer's Supply

The land developer's decision is two-dimensional. He must determine: 1) how many lots to develop (lot quantity subdecision), and 2) the location of the lots (site location subdecision).

The first of these two subdecisions is essentially similar to the production-inventory control decision in a manufacturing firm. In the long run, the number of lots developed must be equal to the number of lots sold; but in the short run, either vacant unsold developed lots or lot shortages may exist.

Because of the large number and part-time nature of many land developers, the dynamic response of land development to long-run demand will, in general, be less stable than in manufacturing. Overoptimistic forecasts of long-run demand lead inevitably to any overdevelopment of land, which later results in sharp contraction of activity to reduce the lot inventory. To simulate the behavior of the land developer's quantity subdecision, the dynamic parameters that produce this unstable pattern must be determined.

The site location subdecision is of a different nature. It resembles in many ways the housing unit type selection subdecision of the household in that a match process between a housing unit type and a land type is involved. The land developer develops a subdivision for a certain class of housing. The nature of the site and its cost (raw land and development costs) are important factors in the match of a housing unit type and a land site. The approach to determining the precise nature of the match will be similar to that previously described for the household-housing unit type combination.

Model Implementation of Land Developer's Lot Quantity Subdecision: The land development lot quantity subdecision is expressed in the model as a rate of development in terms of lots/month. The final land development completion rate in a particular zone (RLDC) is determined from a sequence of six equations which are identical in structure but have different parameters from one zone to another.

The first equation calculates the base land development trial rate as a function of projected sales and lot vacancies. Projected lot sales are the summation of a time-average of past housing sales augmented by a trend correction.

In the formulation, the base land development rate is:

RLDX = PHCC+(VACD-VACL)/TLIA

where:

RLDX - Base land development rate (lots/month)

PHCC - Projected lot sales based on housing constructed (lots/month)

VACD - Lot vacancies, desired (lots)

VACL - Lot vacancies, actual (lots)

where:

VACD = (PHCC)(VACR)

VACL = RLD-HC

and

RLD - Residential land developed (units)

VACR - Vacancy ratio (months)

HC - Total housing units

The term VACD reflects the amount of vacant lots considered normal as a function of the average lot sales rate.

It is necessary to distinguish between unit lots and actual lots because of multi-family dwellings. A unit lot is the number of dwelling units on the lot. If a single-family housing unit is involved, a unit lot and an actual lot are identical; but for multi-unit structures the number of unit lots will depend on the dwelling units on the lot.

The base land development rate (RLDX) is modified by the residential zoning restrictions in such a way that no further land development is permitted after the zoning limit (RLM) has been reached.

RLDY = RLDX if RLD ≤ RLM

RLDY = 0 if RLD > RLM

RLM - Residential zoning maximum (unit lots)

A third equation prevents a negative land development rate should RLDY become negative. Although such a negative rate is possible, as in the case of developed unsold lots reverting back to raw land through plat vacation proceedings in depressed economic times, it was not considered desirable in the model.

RLDZ = RLDY if RLDY ≥ 0

FLDZ = 0 if RLDY < 0

To convert the units of land development rate to actual lots, RLDZ is multiplied by the lot density factor (RDEN).

RLDR = (RLDZ)(RDEN)

RDEN - Residential density conversion factor (lots/unit lot)

Land development occurs over a period of time. To reflect this land development time, a land development time delay is incorporated in the model. Land development completed (RLDC) lags land development started (RLDRS) by a delay period (RDEL).

Residential land available in the zone (RLD) is increased by periodic additions of lots at the land development completion rate (RLDC).

RLD = RLD+(DT)(RLDC)

This completes the relationships for the lot quantity subdecision. It is now necessary to investigate the site location subdecision.

Model Implementation of Land Developer's Site Location Subdecision: Two alternative formulations of the land developer's site location subdecision have been tested.

In the first formulation, site location is completely demand oriented. The above set of equations for the quantity of lots are computed in each areal zone. Site location, therefore, is completely dependent on the demand for lots in each zone. Since this demand may ultimately be traced through the builder back to the household's decision to purchase, the entire model is demand oriented. This version of the site location subdecision is being tested as an alternative formulation; but a second version, also being tested, considers the cost (supply) as well as the demand aspect of the development process.

In the second version of the site location subdecision, the lot development rate (RLDX) is calculated for the entire area being tested. For the Waukesha pilot test, this area included the City of Waukesha and its environs. In general, it would include the area served by a common set of land developers. By calculating demand (RLDX) for the area as a whole, demand is considered as the source for the quantity but not the location of the lots. The lot location is treated instead as a separate decision based on the cost of land development.

In the new site location subdecision, the aggregate lot total in each time period is allocated to individual zonal areas based on a minimization of costs to the developer within the constraints imposed by zoning restrictions. Lot development costs have been determined for five classes of lot sizes based on the type of soil and the physical improvements required. Lot development costs vary significantly with the type of soil. For this reason, a comprehensive soil survey of the kind being conducted in southeastern Wisconsin is essential for simulating the development of new land in this version of this subdecision.

In the model, the aggregate land demand for each lot type in each period is allocated to the model zones using a linear programming subroutine. This subroutine will allocate land using the following relations:

Minimize

$$C_t = \sum_{m=1}^{M} \sum_{n=1}^{N} C_{rmn} R_{mn}$$

Subject to

$$\sum_{m=1}^{M} \sum_{n=1}^{N} R_{mn} + SR_{mn} = R_d + SR_d$$

here:

C_t - Total private land development costs (dollars)

C_{rmn} - Cost of developing a lot of lot type m in zone n (dollars/lot)

R_{mn} - Lots of lot type m in zone n (lots)

S - Service lot ratio (service land, such as retail, school, street land, etc., required to support residential development)

R_d - Total residential land demand (lots)

This alternative approach is based on the hypothesis that land developers will seek out the most profitable locations for lots to satisfy the demand for lots in the area. This is not to say that he will develop the optimal number of lots since his forecast of lot sales is subject to error. The hypothesis implies only that land developers will search for the low cost locations appropriate to the type of lot.

This new formulation of the land developer's site location subdecision in no way implies a change to the behavioral approach to the locational decision of the household. The land developer is a businessman trying to advance his fortunes through land development. His knowledge of land values, development costs, and sales potential is usually highly developed. Although it is recognized that there are many part-time developers who enter and leave the field depending on business conditions, these developers, too, usually possess special knowledge of land. Typically, they work in related fields, such as real estate or insurance or are developing family property. In any case, they usually have an economic orientation since they are developing the land for a profit.

In contrast, the household typically has less knowledge of land values and economic potential. In general, it would seem that the household selects a house and site to satisfy certain housing and locational accessibility needs. As long as these needs can be met at a price considered reasonable, the household makes little attempt to optimize its location economically.

The detailed program equations are listed in Appendix I.

Builder's Decision

In the model formulation, the builder provides housing units in response to household demand. His only decision in the model is a quantity decision, and he affects location only insofar as he provides housing in the areas developed by land developers that are selected by households. The quantity decision of the builder is formulated with the equation structure used for the land developer. The equation parameters will differ, of course, and will vary from zone to zone. Both custom and speculative builders are provided for in the formulation, with the custom builder acting on specific demand and the speculative builder constructing homes for a temporary inventory.

INDUSTRIAL SECTOR FORMULATION

In the endogenous version of the industrial sector, an economic approach, modified by the detailed requirements of particular industries, has been taken to the basic decision of site selection. In essence, this approach is based on the theory that the site selected for a particular firm must possess certain characteristics related to the production and distribution technology of the industry. From the class of sites that comply with these specified characteristics, the firm will then select the lowest cost site available.

Industrial site selection, then, is simulated as a constrained cost minimization process. The decision is functionally similar to the land developer's decision in the residential sector in its cost minimization approach. It differs, however, in the more significant role played by technological constraints. The requirements for an industrial site are likely to be more numerous and more carefully analyzed than the requirements of a residential housing site.

The linear programming formulation of the industrial sector would take the following form:

Minimize $$C_t = \sum_{i=1}^{J} \sum_{n=1}^{N} C_{in} I_{in}$$

Subject to $$\sum_{i=1}^{J} \sum_{n=1}^{N} I_{in} + SI_{in} = I_{id}$$

$$\sum_{i=1}^{J} I_{in} \leq I_{nz}$$

$$I_{in} \geq 0 \text{ if } I_{in} \epsilon J_i$$

or $$I_{in} = 0 \text{ if } I_{in} \epsilon \overline{J_i}$$

Where:

C_t - Total industrial land development costs (dollars)

I_{in} - Industrial land for industry i developed in zone n (acres)

C_{in} - Cost of developing industrial land for industry i in zone n (dollars/acre)

I_{id} - Total regional demand for industrial land in industry i (acres)

S - Service ratio--ratio of service land area to industrial land area

J_i - Set of land meeting requirements for industry i

$\overline{J_i}$ - Set of land not meeting requirements for industry i

I_{nz} - Capacity limit for industrial land in zone n

Costs (C_t) of industrial land development are minimized subject to the restrictions that all the land required for each industry (I_{id}) must be satisfied and no firms may be located in an area that does not satisfy the minimal requirements for that industry.

During each time period, total industrial land demand will be calculated, based on the number of regional firms originating or moving within the Region and the new firms entering the Region. These firms will be located to particular areas by the linear programming subroutine.

The primary data requirements for the industrial sector of the model, then, are the site selection criteria for each industry or group of industries and the land development costs. The cost data are being collected for use in the Land Use Plan Design Model and for general land use planning purposes. Site selection criteria will be based on the Stefaniak study [15] recently conducted in the Milwaukee area. In this study site selection criteria were obtained for 759 plants representing all manufacturing industries in the Milwaukee area. This information is available in punched card form and so is in a form suitable for immediate analysis.

Use of the site selection criteria will require the separation of "necessary" criteria from "desirable" criteria since the basic industrial land allocation concept considers criteria from an "all or nothing" point of view. In the industrial site selection decision, all sites not complying with the required characteristics are eliminated from consideration; and cost minimization takes place within the acceptable site area. The most difficult analytical task will be the separation of "desirable" from "required" site characteristics.

SERVICE, SPECIAL, AND AGRICULTURAL SECTORS

Land uses in the service sector, such as local retail trade and services, streets, and other categories, will be allocated based on service ratios required to support primary residential and industrial land uses. These service ratios for test simulation of past land use development will be based on analysis of historical service ratios. The interacting nature of the residential and service sectors should be emphasized. Service land use depends on residential and industrial land use, but an increase in service land use also influences further residential land use through the accessibility effect. Future service land use plans will be based on service design standards.

Land uses in the special sector will be based on programmed inputs since these land uses are usually based on project-type decisions by government or the private sector of the economy. In fact, one of the primary areas of interest in model simulation will be the effect of freeways and other elements of the transportation plan on land use development.

The agricultural sector in historical land use simulation will be a residual land use, in that land previously in agriculture will be transferred to residential, industrial, or associated

[15] *Norbert J. Stefaniak, Industrial Location Within the Urban Area; University of Wisconsin, Madison, 1962.*

service land uses. In future land use plans, however, an attempt will be made to preserve certain agricultural lands in the land use plan design based on their productivity.

ESTIMATION OF MODEL PARAMETERS

A combination sampling and regression approach was used to estimate the parameters of the Land Use Simulation Model.

Most of the household parameters are based on data collected through the household history portion of the home interview origin and destination survey. This household history survey detailed the home and work locations, together with other data on the household and housing characteristics of the sampled households, for the period 1950-1963. With this data, it was possible to classify the sampled households into type clusters and to determine their parameters for the turnover, housing preference, and site location decision formulations.

With the sampled households classified into types, parameters were estimated from the average values of historical decision patterns. The primary problem was the classification process itself. To accomplish this classification, a special set classification program was developed that would decompose a household set into subsets with common characteristics. Examples of household characteristics are the age, income, and education of the head of the household. Each subset of each characteristic is designated an attribute. Examples of attributes for the characteristics just described are: age of head of household under 35, income over $10,000, and college education.

The inputs to the set classification program are all of the selected characteristic-attributes of each sampled household. The output is a set of household types grouped according to common attributes. The minimum size of the smallest type subset is determined by the user. This size should depend on the size of sample being classified.

Parameters relating to the lot and housing quantity subdecisions of the land developer and builder were determined by regression analysis. Using a current land use inventory and historical records of subdivision plats and building permits, it was possible to synthesize a history of land development and construction beginning in 1950. From this history the dynamic parameters of these subdecisions were calculated with a regression analysis.

The site location subdecision of the land developer depends primarily on the relative costs of raw land and residential development. These costs depend on the topographical and soil characteristics of the area. Data from the regional soil survey provided the base for costs by zonal area. Detailed engineering cost estimates were developed for land development of varying lot sizes on three suitability classes of soils.

In the service sector of the model, the service ratios for auxiliary land use were determined, either by historical ratios or planning design standards, depending on the model application. Historical ratios are suitable for forecasts of uncontrolled land development, and design standards are preferable for plan implementation.

Industrial sector land costs for the endogenous version of this sector were also based on the soil survey. There was a significant amount of common data used by the Land Use Plan Design Model and the Land Use Simulation Model. Both used the land development cost data, and both required service ratios for auxiliary land uses.

The land requirements of the various industries in the industrial sector were determined by a special study made by Professor Norbert J. Stefaniak of the University of Wisconsin-Milwaukee.[16] This study provided a comprehensive analysis of the land and public facility

[16] *Ibid.*

requirements of industries in the Milwaukee area. Although this study was prepared independently of the planning program of the Southeastern Wisconsin Regional Planning Commission and had no connection with the development of the model as such, it has proven to be of critical value to the industrial sector of the model.

MODEL PILOT TESTS

The same 1950-1962 period of historical land development in the City of Waukesha and its environs, previously used for the test of the Land Use Plan Design Model, was also used for the tests of the Land Use Simulation Model. Only the residential sector and its associated service land uses were simulated in initial tests since the Waukesha area was too small for a realistic industrial land use simulation test. In the simulation tests, the measure of model effectiveness was quite different from that for the design model. Instead of developing a plan design consistent with a given set of objectives, costs, and constraints, the simulation model was expected to "act like" the actual land development process over time.

In the program of model tests, two types of simulation were performed. In the first type, the actual land-housing demand was provided as the land developer's forecast thereby bypassing the household as a synthetic demand generator and the land developer as a forecaster. The results of this test are illustrated in Figure 4. The purpose of this artificial type of simulation was to test the accuracy of the land developer's decision simulation with a perfect demand forecast input.

The second type of simulation, the results of which are illustrated in Figure 5, involved the complete model with an internal generation and forecasting of land-housing demand. This two-part simulation permitted a separation of model inaccuracies caused by incorrect household demand from those of the land developer's site selection decision.

The initial test results were quite encouraging. The perfect demand simulation, as illustrated in Figure 4, provided a model accuracy performance of 90.3 percent. Stated another way, 90.3 percent of the actual land development was accounted for in the land use simulation. From Figure 4, it will be noted that the accuracy of the model varied somewhat from zone to zone. Some zones were perfect, while others varied significantly from the actual development. In general, the zones with the more extensive development in terms of the number of lots were simulated more accurately, while those with fewer lots were less accurately portrayed. Such a result is understandable in any simulation since the statistical law of large numbers results in an averaging of random errors in the simulation. This same effect occurs in traffic assignment simulation models where links of low volume are less accurate than links of high volume.

Evaluated on a statistical basis, the zones with perfect results would seem difficult to explain, but these results are easily understood in the light of the constrained optimization nature of the land developer's decision. The use of a recursive programming algorithm for this decision simulation means that some zones will be developed to the limit of some physical or legal constraint. This is precisely what happened in the zones in question.

The measure of model accuracy was obtained by dividing the summation of the absolute value of the errors in each of the zones by the total number of lots actually developed. A similar result may be calculated by dividing the average absolute value of the error in each zone by the average size of each zone. Another measure of model accuracy is the statistical correlation coefficient, which was 0.97 for this simulation test run.

The perfect forecast simulation run is really a test of the basic theory of the model relating to the land development process. This theory, as previously discussed, states that a land developer will seek out the lowest cost areas within the physical and legal environment in which he must operate (land developer's site selection subdecision). While more extensive tests of this theory in a number of regional areas will be required before it can be accepted universally, the theory has found some support at least in its application to southeastern Wisconsin.

A second aspect of the model and its supporting theory was tested in a second series of runs of the complete model in which the demand for land and housing was internally generated within the model. The land developer bases his plans for development on forecasts of land demand generated within the model. Model inaccuracies are now increased by the errors in the model simulation of the land-housing demand for the area. These new errors reflect primarily the accuracy of the data used to determine the household-housing-land demand generator used in the model. Parameter errors in the land developer's lot quantity subdecision will generally only influence the timing rather than the total amount or the size distribution of the lots developed.

Comparative results indicate a total model accuracy of 80.5 percent and a correlation coefficient of 0.91. The errors introduced into the model by the internal generation of demand provide a comparison of the error effects of the site location of land supply as opposed to the total quantity and lot size distribution of land supply. Test results would indicate that total supply, as well as site location, are both key factors in the effectiveness of a land use simulation model. The effects of land demand forecast errors were somewhat amplified by the model computational procedure, which was affected by time-phase errors in forecasts. This sensitivity to phase has been removed in the new computational procedure.

The validity of the above test results can be appreciated only if the source of the data used to estimate the model parameters is clearly understood. Most land use models previously developed have used some form of regression analysis to estimate the model parameters. Such regression analysis involves the use of time histories of the variables being simulated. Because historical data is used to estimate the model parameters, this same history cannot be used to validate the model. The parameters of the Land Use Simulation Model were not estimated from regression analyses of time series but from independent estimates based on surveys, such as the household history of sampled households or the engineering estimates of costs based on the type of soil. This independence between the time series history of the model variables and the samples used to estimate the model parameters provides a sound basis for model validation tests.

Leaving the world of statistics for a moment, it is important to consider the practical usefulness of the Land Use Simulation Model in relation to the land use-transportation planning process. Although the simulation model might be employed for a variety of purposes, three primary functions predominate in the current regional planning program:

1) Forecasting the future land use pattern based on a given aggregate forecast of population and the continuance of existing and committed public works programs and land use control policies.

2) Testing the feasibility of land use plans and providing a vehicle for the experimental design of land use control policies and public works programs to achieve these plans.

3) Providing the land use pattern (forecast or plan) for the determination of spatially distributed travel demand (trip generation) which serves as the input to the transportation planning process.

Each of the above functional requirements involves a determination of the structure of the land use pattern. Pilot test results would seem to indicate that the simulation of the land use pattern will probably be more accurate than both the input and the output of land use simulation. The input forecast of 1990 population will do well to achieve an accuracy of 90 percent. In fact, it is quite likely that only a continual monitoring and adjustment of this forecast will provide the accuracy necessary for the continuous planning process.

The output of land use simulation (from the viewpoint of transportation planning) is trip generation. Experience to date in trip generation analysis indicates significant difficulties in developing accurate relationships from current data much less from forecasts of future activity. The accuracy of the simulated land use pattern should not seriously restrict future trip generation forecasts.

Given the goals of the regional planning program in southeastern Wisconsin, the predominance of the second of the above three functions, land use plan test, is obvious; and it is in this function that the model is at its best. Experience to date with simulation models in both industry and government have confirmed the usefulness of such models in policy design. Test results of the Land Use Simulation Model reinforce this general conclusion in land use planning applications.

FUTURE MODEL DEVELOPMENT AND APPLICATION PROGRAM

The model application schedule now calls for simulation of land development at a regional level in the coming months. Such a regional simulation will be used to quantitatively test the implementation feasibility of the land use plans synthesized in the land use-transportation study. Primary preparations for such tests will relate to the estimation of model parameters from household, land use, soil, utility, and other survey data.

Concurrent with this application of a larger-scale version of the Waukesha model, however, a number of extensions and improvements to the model will be attempted. These will include:

1) Further extension of the recursive programming concept in the land developer's decision to provide for behavioral zonal capacity restrictions that depend on the previous status and rate of land development in the zone. This change may be very effective in improving model accuracy performance since the lack of such adjustment seemed to be the primary source of error in the Waukesha model.

2) A more efficient computational procedure has been developed, which will significantly reduce the computer time needed for a simulation run. In this new procedure, a land use demand forecast submodel will generate future annual land use demand by residential density class (low, medium, and high). This series of annual forecasts will then serve as the input to the recursive programming submodel which will distribute this land demand spatially in the Region. Model running times less than half that required in the previous approach are estimated. This reduction in running time has been accomplished without any sacrifice in the accurate representation of the theory underlying the model. The new procedure, in fact, adds to the flexibility of the model and removes the sensitivity to time-phasing forecast errors described in the previous section.

3) Attempts will be made to improve the accuracy of internal land demand generation by a reclassification of household types.

4) Regional retail trade and service land uses will be simulated concurrently with residential land use. Previously, only local retail and service land uses were simulated as supporting residential land uses.

5) Industrial land use will be simulated on an experimental basis concurrently with residential and associated land uses.

6) Residential, commercial, and industrial land use spatial allocation will interact through the medium of accessibility constraints in the regional version of the model.

Each regional shopping center and industrial employment center will be limited in the population it can support within its area of service. These accessibility constraints will provide for simultaneous consideration of site characteristics and accessibility in model operation.

Appendix II

Table 2A

LAND USE SIMULATION MODEL HOUSEHOLD TYPE CLASSIFICATION

Household Types					
Type	Education*	Occupation*	Income*	Age*	Sex-Race
1	College	White Collar	$ 0 - $ 8,000	Under 35	Male-White
2	College	White Collar	0 - 8,000	Over 35	Male-White
3	College	White Collar	8,000 - 12,000	All	Male-White
4	College	White Collar	12,000 - 20,000	All	Male-White
5	Non-College	White Collar	0 - 4,000	All	Male-White
6	Non-College	White Collar	4,000 - 8,000	Under 35	Male-White
7	Non-College	White Collar	4,000 - 8,000	Over 35	Male-White
8	Non-College	White Collar	8,000 - 12,000	All	Male-White
9	Non-College		12,000 - 20,000	All	Male-White
10		Blue Collar	0 - 4,000	All	Male-White
11		Blue Collar	4,000 - 8,000	Under 35	Male-White
12		Blue Collar	4,000 - 8,000	Over 35	Male-White
13		Blue Collar	8,000 - 16,000	Under 35	Male-White
14		Blue Collar	8,000 - 16,000	Over 35	Male-White
15	All	All	All	All	Female-White
16	Miscellaneous				Female-Non-White

**All household type characteristics are based on the head of the household.*

Appendix IIA

Table 3A

LAND USE SIMULATION MODEL HOUSING UNIT TYPE CLASSIFICATION

Housing Unit Types			
Type	Structure	Status	Market Value
1	Single Family	Own	$ 0 - $10,000
2	Single Family	Own	10,000 - 15,000
2	Single Family	Rent	0 - 60
3	Single Family	Own	15,000 - 20,000
3	Single Family	Rent	60 - 100
4	Single Family	Own	20,000 - 25,000
4	Single Family	Rent	100 - 150
5	Single Family	Own	25,000 & Over
5	Single Family	Rent	Over $150
6	2-3-4 Family	Own	0 - 15,000
6	2-3-4 Family	Rent	0 - 60
7	2-3-4 Family	Own	15,000 - 25,000
7	2-3-4 Family	Rent	60 - 100/mo.
8	2-3-4 Family	Own	25,000 & Over
8	2-3-4 Family	Rent	100 - 150/mo.
9	5-19 Family	Rent	60 - 150/mo.
10	Miscellaneous		

LAND USE SIMULATION MODEL INPUT DATA

The following input data in the order listed is required for the operation of the Land Use Simulation Model:

1. TIME (Variable, Initial Condition)
The initial starting time is loaded. It is not always 0.0, which represents the base year 1950, because some runs were begun in a later year.

2. DT (Parameter)
This is the recursive time interval used in the model.

3. ENDTIM (Parameter)
This time represents the period in which the simulation run will end.

4. L (Parameter)
Number of lot types.

5. M (Parameter)
Number of housing unit types.

6. N (Parameter)
Number of household types.

7. PRTPER (Output Parameter)
The time interval between printed outputs of simulation results.

8. ACCPER (Parameter)
The time interval between accessibility calculations.

9. NZ (Parameter)
The number of zones.

10. DU (Parameter)
The dwelling units in each building for each housing type.

11. TO (Parameter)
The turnover rate (moves/year) for each household type.

12. HVR (Parameter)
Housing vacancy ratio, which determines the desired number of unsold developed lots allowed to exist without reduction of current land development.

13. THIA (Parameter)
Housing inventory adjustment time, which determines how rapidly builders increase or decrease building activity to adjust for vacancies.

14. TLIA (Parameter)
Same as 13 above as applied to land development rather than building.

15. TS, HIT, HPT, TSH, HITH, HPTH (Parameter)
Parameters for exponential smoothing or forecasting to simulate the forecasting behavior of land developers and builders.

16. TDEL (Parameter)
Household relocation time represents the delay in the relocation of a household. So far, it has been used only for programming convenience to maintain an intransit level in the central subsector.

17. HN (Exogenous Variable)
New household formations and in-migrants for each household type are both included in HN which represents the primary exogenous variable input to the model.

18. TCX (Parameter Matrix)
This matrix determines the transition of housing from one type to another as a function of time.

19. TRX (Parameter Matrix)
This matrix determines the transition of households from one type to another as a function of time.

20. CP (Parameter Matrix)
This matrix transforms the increase in the number of new households (formations and in-migrants) into housing demand by housing type.

21. PC (Parameter Matrix)
This matrix transforms the available housing by type in each zone into a housing preference factor for each household type which determines the distribution of relocating and new households to each zone.

22. HLX (Parameter Matrix)
This matrix transforms housing demand by type into land demand by lot type.

23. HLY (Parameter Matrix)
This matrix transforms the available land by lot type in each zone into the land available for each housing type for new construction.

24. OMC (Parameter)
Some of the relocating households move out of the region. The out-migration constant determines