Chapter 3: Overview of Computer Simulation Modeling Approaches and Methods

Robert E. Manning
Robert M. Itami
David N. Cole
Randy Gimblett

Capturing Behavior of the System

The field of simulation modeling has grown greatly with recent advances in computer hardware and software. Much of this work has involved large scientific and industrial applications for which substantial financial resources are available. However, advances in object-oriented programming and simulation methodology, concurrent with dramatic increases in computer capabilities and reductions in computer hardware costs, have meant that the benefits of simulation can be extended to areas that previously have been impractical. This includes recreation management.

The challenge of simulation modeling is to capture the essential behavior of the system being modeled. In outdoor recreation, this means capturing and representing the characteristics of the physical environment (for example, a system of trails, roads, waterways, and/or facilities) and modeling the behavior of visitors as they interact with the environment and with each other. In the most basic sense, models have three components: (1) input variables that describe the system being modeled, (2) software and associated modeling approaches designed to process these input variables, and (3) output variables that are useful to planners, managers, and scientists. This chapter outlines these components for recent modeling efforts in park and wilderness management. Chapter 4 provides a series of case studies designed to illustrate these basic components of simulation modeling and demonstrate their potential usefulness.

Modeling Approaches and Software

Simulation modeling software is needed to process input variables, generate data analyses, and produce output. Three approaches to modeling and simulation of relevance to recreation travel simulations are trace, probabilistic, and rule-based agent models.

Trace, Probabilistic, and Rule-Based Agent Simulations

Trace simulations directly simulate travel itineraries collected in the field. Visitor arrival, trip itineraries, and duration of stay at destinations are simulated directly from survey data rather than using probability distributions or random numbers. These simulations are useful for examining existing pattern of use, and are often used to validate probabilistic and rule-based simulation models that are derived from the same data. Probabilistic simulation models are based on a representative sample of visitor trip itineraries. Visitors' trips are then modeled based on the probability of a visitor selecting a single trip itinerary out of the entire sample, or alternatively, the probability of selecting the next destination based on the probability distribution of all destinations originating from the current destination. Probability models are the standard method for modeling baseline conditions. Probability distributions for either trip itineraries or origin-destination pairs are a convenient way to "ramp up" numbers of visitors (increase visitor use levels) in a simulation, since a standard assumption is that as the number of visitors increase, the distribution of trip itineraries will remain the same.

Probabilistic simulation assumes that the distribution of trip itineraries in the future will remain similar to the distribution today, regardless of how the system changes. This may be an inappropriate assumption for a system that is changing dramatically. Consequently, probabilistic simulation may not be an appropriate way to model behavior in new recreation settings or in existing settings where management policies may introduce new travel networks, delete existing travel networks, or where behavior may change due to changes in recreation mode or mix of recreation types. For these situations, rule-based simulation may be more appropriate (Itami and others 2004).

Rule-based simulations use autonomous agents. The agents are autonomous because once they are programmed, they can move about their environment, gathering information and using it to make decisions and alter their behavior according to specific environmental circumstances generated by the simulation. Each individual agent has its own physical mobility, sensory, and cognitive capabilities. Because autonomous agents have their own reasoning system for navigating a travel network, the travel network must be attributed with properties to which the agents respond. These attributes may be in the form of attractions such as scenic views, interpretive centers, picnic areas, or playfields, and detractors such as hazardous areas, extreme weather events, or other environmental factors that would constrain movement or cause visitors to avoid an area. It is these attributes and the attributes of other agents that determine agent rules.

Terminating and Steady-State Simulations

A second important choice in simulation modeling approaches is whether to design simulations to be terminating or nonterminating (steady-state). Terminating simulations model events that have a specified length, while a steady-state simulation models situations in which there is no natural event to specify the length of a simulation run (Law and Kelton 2000). The choice between these two should be made on the basis of the situation being modeled and the desired modeling outputs. A terminating simulation has a known initial state (usually zero) and a known ending state. For day use issues, it is clearly appropriate to use terminating simulations to describe what happens over a given day, based on data representing the typical arrival sequence for a day.

When the situation of interest involves people on multiday trips, modeling individual days makes little sense. Nor does it make sense to model the entire year or season of use. What we are usually interested in understanding is how the system operates when at its full operating level (often at its maximum or peak). This type of situation might be modeled using a steady-state simulation. A simulation is called steady-state because the simulation, after an initial “warm up” period, is designed to replicate system behavior over the long run at a given level of production or capacity.

It is currently unclear whether it is more appropriate to model multiday backpacking trips using terminating or steady-state simulations (see for example the John Muir Wilderness case study in Chapter 4). This situation has some characteristics that seem best handled with steady-state simulations and some that seem best handled with terminating simulations. Regardless of appropriateness, steady-state simulations are more challenging to conduct and analyze. They must be run over long periods to get a reliable average measure of system behavior that is not biased by short-term effects of random variables and auto-correlation. The results of steady-state simulations must be carefully interpreted since they can overestimate parameters if the actual duration of steady-state conditions in the field is relatively short in comparison to the run-lengths required to get valid simulation results (Law and Kelton 2000). Currently, we do not have a good understanding of how to use steady-state simulations in recreational contexts.

General Purpose Simulation Software and Special Purpose Simulators

Commercially available general purpose simulation software packages are usually developed with business, industry, and government applications in mind. However, it is possible to use this general software to model outdoor recreation behavior. For example, several of the case studies described in the next chapter have adapted the simulation software, Extend, developed by Imagine That, Inc., to recreational applications. Special purpose simulators, however, are developed specifically to handle specialized applications. Several of the case studies described in the next chapter have used RBSim, developed by GeoDimensions Pty Ltd. This is a special-purpose simulator designed to build simulations of recreation behavior on linear networks. Special-purpose simulators will have more automated features specific to the application of concern. General-purpose simulation software can also be modified to include automated features specific to the application of concern (modeling outdoor recreation behavior).

Model Inputs

Simulation models require several types of input data that can be obtained from several sources. Principal types of input include data on the travel network, the environment, visitor characteristics, and, in some cases, decision rules.

Travel Network

In all of the applications of simulation modeling to park and wilderness management to date, recreation use is constrained to linear travel networks. The travel network may be represented by a road or trail system, river, or flight path. Travel networks are described as a series of connected links that are terminated or joined by nodes. In lay terms, links are referred to as trail, road, or river segments. Nodes are points that terminate a link or join three or more links at an intersection, or may be points that mark a destination along a single link (for example, a campsite or an attraction site). Travel networks are complete if there is a path from any node in the network to any other node in the network. All simulations described in Chapter 4 require complete travel networks.

Links and nodes can have properties or attributes assigned to them. For links, associated attributes might include name, maximum travel speed, average travel time, number of travel lanes, surface type and condition, steepness, length, and width. Node attributes might include name, a list of one or more facilities, visitor capacity, or average visit duration.

Travel networks can be derived from existing maps, or collected using traditional land survey techniques, downloaded from Geographic Information Systems (GIS), or collected using Global Positioning Systems (GPS).

Environmental Data

Environmental data may be required for specific models depending on the nature of the simulation. For example, a Digital Elevation Model (DEM) may be required if visual encounters are recorded or if pedestrian speeds are to be affected by uphill and downhill travel directions. Visual encounters occur when two groups see each other only from a distance. The DEM is represented by an evenly spaced grid of elevations. It is normally imported from a GIS system in a standard exchange format.

If simulation is used for facilities management (for example, to size parking lots, campgrounds, viewing platforms, or other facilities), additional information on each facility must also be collected, including the location and capacity of the facility and the typical duration of stay. Also, if queuing behavior at a facility is to be simulated, information about the service times and maximum queue length might also be required.

Visitor Characteristics

In its simplest form, for probabilistic models, a simulation model of recreation behavior requires information about travel mode characteristics (for example, foot, car, bus, or horse), travel speed, and a trip itinerary. In all cases, data collected must be in the form of a census or a representative sample. The sampling period must be appropriate to the needs of the simulation. For terminating simulations, the sample should be over the complete day or other period of interest. For steady-state simulations of peak use, sampling should be done during the peak period of use.

Visitor characteristics are generally collected using either direct observation or survey techniques. The trip itinerary is a list of destinations and visit durations (or "delays") at destinations. The trip itinerary will usually have the following:

- An entry node (trailhead, park entry, and so on)
- A series of one or more destination nodes.
- An exit node (may be the same as the entry node for round trips).
- Arrival time (the date and/or time the visitor arrives). Often the arrival time is represented by an arrival curve or arrival sequence in which the number of visitors arriving per hour or day is provided.
- Visit duration (or "delays") at nodes. This may be represented by a mean duration or a statistical distribution.
- Overall trip duration.

In a single simulation there may be many different itineraries (as many as one itinerary per visitor) or itineraries may be generalized as a set of probabilities for moving from one destination to another. Probabilistic itineraries are generated using statistical analysis of groups of itineraries, with probabilities generated from one destination to another.

Travel itineraries can be obtained from reservation systems where trip itineraries are specified (such as backcountry permits), or can be collected from trip diaries, GPS tracking, race timing equipment, video monitoring, or self-administered trip recording. Chapter 4 includes a case study in which waypoint signs were used to establish travel itineraries for a complex frontcountry trail system at Mount Rainier National Park.

Arrival times can be collected from traffic counters, survey data, pedestrian counting systems such as turnstiles, pressure-sensitive pads, or infrared counters. Delay times at facilities and destinations can be obtained by trip diaries, onsite observation, video monitoring, GPS tracking, race timing equipment, or directional pedestrian counters.

Decision Rules

Decision rules are required for rule-based simulation models. In this type of simulation, trip itineraries may not be known because the travel network may not yet be developed or because management conditions have changed (such as trail or road closures), which alter existing itineraries. Decision rules may also be required where the behavior or itinerary of a visitor may change based on conditions that are generated through the course of the simulation, such as parking areas becoming full, crowded conditions at visitation sites, weather changes, or other events that may alter onsite behavior.

Agent rules are a set of user-defined behaviors using a stimulus/response or event/action framework. Rules are determined by observing onsite behavior, surveys or interviews, or "walkthroughs" whereby visitors or people familiar with the pattern of use are asked to systematically trace a trip on a map and identify decision points, destinations and attractions, site detractors, or other visitor- or site-related conditions that would alter path selection. This process is repeated for each visitor type or itinerary. Generally the rules must be tested, by simulating one agent at a time, to fine-tune the rule conditions, the sequence of rules, and the complexity of rules needed to achieve the desired path selection behavior. A rule of thumb is to keep rules as few and simple as possible.

Boolean logic can be used to combine two or more stimuli to create complex conditions for behavior. Broolean logic organizes concepts in sets that are controlled by the operators OR, AND, and NOT. An example of a rule is:

If (TravelMode = 'Car' AND Locale = 'Twelve Apostles' AND LocaleEntry = True) THEN Find Carpark

In this example, the rule would only apply to agents arriving in a car at the entry to the Twelve Apostles Locale. If these three conditions are met, then the agent is directed to find a carpark. This directive then triggers the agent to execute its complex wayfinding logic to create a new itinerary.

It is also possible to have probabilistic rules. For example, we may know from count data that 60 percent of all visitors stop at the visitors center. We could construct a probabilistic rule by assigning the rule a probability of 0.60. When the agent triggers a probabilistic rule, it generates a random number in the range of 0 to 1. If the number generated is less than or equal to the probability assigned to the rule, then the rule is executed. If it is greater than the probability assigned to the rule, then the rule is ignored. Refer to the Port Campbell National Park (Twelve Apostles) case study for more detailed information on rule-based agent modeling.

Model Outputs

Simulation models can generate a great variety of output variables. These variables should be specified prior to model design and development, and should be formulated on the basis of their potential usefulness to park and wilderness managers. Commonly used output variables include use density, encounter, and queuing time measures. Use density measures report the number of visitors related to space or time. For example, a simulation model of a backcountry trail and campsite network could report the number of visitors (hikers, bikers, horseback riders) that traverse each trail segment per day or the number of visitor groups at each campsite per night. Other related output variables include people-at-one-time (PAOT) at attraction sites and people-per-viewscape (PPV) along heavily used trails. A simulation model of a trail and campsite network can also report the number of times visitor groups encounter one another along trails or at campsites. Encounters can be recorded by type of visitor (hikers, bikers, horseback riders), by type of encounter (meeting, overtaking, visual), by place, and by unit of time. Encounter estimates are a particularly important output. Because encounter type and number can influence visitor experience, standards are often written for encounters, which are difficult to directly monitor. Finally, models can output queuing or waiting times for visitor facilities or services. A variety of output variables are included in the case studies described in Chapter 4.

Since models are driven by random samples from probability distributions, the output from two different simulation runs can be quite divergent. Consequently, it is important to incorporate replication into simulation modeling. Replication can involve either running many different simulations or, for steady-state simulations, running a single long simulation that is divided into "batches" that serve as replicates (Law and Kelton 2000). In either case, outputs should be reported as means with confidence intervals.

Model Validation

An oft-neglected step in the model-building process is validation of the model. Validation is the process of making certain that the simulation model provides an accurate representation of the system being modeled. As Law and Kelton (2000) note, if the model is valid, decisions made using the model would be similar to those that would be made if it were possible to physically experiment with the system. There are at least three important steps in model validation (Law and Kelton 2000; Naylor and Finger 1967; Schechter and Lucas 1978). First, the model should be checked for face validity. That is, it should work in ways that seem reasonable to those who know how the system should operate. Second, the operating assumptions of the model should be empirically tested. Sensitivity analyses can be used to make certain that outputs change in predicted ways when important model variables are changed. Finally, it is important to assess how closely model output data resemble those that might be obtained from field data.

Schechter and Lucas (1978) discuss various ways they assessed the validity of the Wilderness Use Simulation Model. A common validation technique involves comparing model output to field observations. Simulation models are normally developed initially to describe current or "baseline" conditions in a park or wilderness area. For example, input variables on visitor use levels and patterns for a trail and campsite network might be used to develop a model designed to estimate trail and campsite encounters. Field observations of trail and campsite encounters might be collected through direct observation or self-reports of visitors, and these data compared to model estimates to test the degree to which model estimates are accurate or "valid."

In the case studies reported in Chapter 4, face validity has generally been assessed and sensitivity analyses have been conducted. However, comparisons of model output to field data have either not been conducted or have been based on limited field data. Therefore, our confidence in the validity of models is limited, as is our knowledge of the precision of results. Where limited field data were collected, it is not possible to conduct a more rigorous validation than to simply compare observations with outputs and conclude whether they seem similar enough to have confidence in the model. With a sufficient sample of field observations it is possible to rigorously estimate the accuracy of the model. Freese (1960) shows how standard chi-square tests can be used to estimate the accuracy of the model.

Using the Simulator

As noted above, the simulator can be used to describe the existing spatio-temporal distribution of use. Output from the simulation can be valuable to management, given the difficulty of obtaining such information in any other way. The simulator can also be used to monitor crowding-related indicators, such as number of encounters, persons-at-one-time or persons-per-viewscape, either to describe the current situation or to determine whether standards for such indicators are being violated. The simulator provides a much more cost-efficient way to gather such data than monitoring these indicators directly.

The simulator can also be used for predictive purposes. For example, it can be used to predict the maximum amount of use that can be sustained without exceeding some crowding-related standard, such as trail encounters per day, by running scenarios with various use levels until the use level that barely complies with the standard is identified. A wide variety of alternative management scenarios can be simulated to predict the outcome of such actions as changes in the timing and distribution of use, changes in the travel network, and changes in facility type, location, or capacity. Alternatives are evaluated by running the simulator after making changes in visitor data or travel network data. Outputs from several different scenarios can be compared to identify those most closely aligned with desired future conditions. Many of these uses of simulation are illustrated in the case studies in Chapter 4.

References

Freese, F. 1960. Testing accuracy. Forest Science. 6: 139–145.

Itami, R.; Raulings, R.; MacLaren, G.; Hirst, K.; Gimblett, R.; Zanon, D.; Chladek, P. 2004. RBSim 2: simulating the complex interactions between human movement and the outdoor recreation environment. Journal of Nature Conservation. 11: 278–286.

Law, A. M.; Kelton, W. D. 2000. Simulation modelling and analysis 3d ed. Boston, MA: McGraw-Hill. 760 p.

Naylor, T. H.; Finger, J. M. 1967. Verification of computer simulation models. Management Science. 14: 92–106.

Shechter, M.; Lucas, R. C. 1978. Simulation of recreational use for park and wilderness management. Baltimore, MD: Johns Hopkins University Press. 220 p.