PURCHASED BY USDA FOREST SERVICE
FOR OFFICIAL USE ONLY

INTERFACES
Vol. 10, No. 5, October 1980

Copyright © 1980, The Institute of Management Sciences
0092-2102/80/1005/0011$01.25

VALIDATING A LARGE SCALE SIMULATION MODEL OF WILDERNESS RECREATIONAL TRAVEL*

Mordechai Shechter

Department of Economics, University of Haifa, Israel

and

Robert C. Lucas

USDA Forest Service, Forestry Sciences Laboratory, Drawer G, Missoula, Montana 59801

ABSTRACT. A large-scale simulation model of the use of outdoor recreation areas, especially ones with dispersed recreation patterns, has been developed that provides a means for experimenting with modifications of use or area conditions to determine effects on use patterns and congestion. The model and its applications to date, especially to the Desolation Wilderness in California, are briefly discussed.

The main purpose of this paper, however, is to illustrate a validation procedure for such a large-scale simulation model, employing a series of validity tests. Because the question of validation has always constituted a thorny problem in the field of simulation, the paper attempts to show — in the context of a case study — how tests can go beyond the customary "tests of reasonableness" often employed in such cases.

The question of the validity of simulation models is often raised and discussed in the literature on simulation. The problem is simply this: To what extent can a manager trust the results of a simulation model as a guide in solving problems in the real system?

Unfortunately, there is no universal agreement on the best method(s) to test validity of a model, whether analytic or simulation. Hermann [1967] has suggested five different tests for models whose predictive powers have not yet been tested. Only one truly qualifies as a scientific, objective test. It measures the degree of correspondence between the model's prediction of observable events and the events themselves when they occur (under a policy which was simulated beforehand). If a model cannot be shown to be highly reliable in terms of construction and predictive ability, and if the prospective users of the model cannot be convinced of its reliability as a management tool, then the researcher's efforts have been fruitless.

Our purpose in this paper is to illustrate an approach employing a series of reasonable tests — not merely a test of reasonableness — to check the validity of a

SIMULATION; RECREATION

*This research was carried out under a cooperative agreement between Resources for the Future, Inc. and the U.S. Forest Service, which provided funds for the project. The first author was a visiting scholar at Resources for the Future when he was associated with this project. Portions of this paper are based on the book, *Simulation of Recreational Use for Park and Wilderness Management*, by the authors, published by Johns Hopkins University Press for Resources for the Future.

Aldo Leopold Wilderness Research Institute: Publication # 76
CITATION : Shechter, Mordechai; Lucas, Robert C. 1980. Validating a large-scale simulation model of wilderness recreational travel. Interfaces. 10(5): 11-18.

large-scale simulation model which is now being used by public land managers. Our basic aim has been to demonstrate a method to overcome the problem of validation of such a model. We did not aim to present a transferable technique, necessarily; other models and situations might require another mix of strategies. We describe first the model itself and the real-life problems it is designed to solve.

WILDERNESS RECREATION MANAGEMENT PROBLEMS

The number of people visiting most kinds of outdoor recreation areas has been growing as long as records have been kept. This growth often creates problems for management, with the nature of the problems depending on the type of area and the management objectives established for it. In the United States, growth in use of dispersed recreation areas has been rapid [Lloyd and Fischer, 1972], and the resulting problems of congestion and resource damage have been difficult for managers to solve. This is especially true in wilderness areas which, by law, are supposed to contain substantially unmodified, natural ecosystems and provide opportunities for solitude.

Studies of the distribution of wilderness recreational use show very uneven use patterns; they are heavily concentrated on certain fairly small portions of each area, while larger portions receive little use. Similarly, summer weekends usually experience sharp peaks in use. Redistribution of use in time and space seems to offer considerable hope for reducing the adverse effects of heavy use [Hendee, Stankey, and Lucas, 1978].

In addition to the legal objectives for wilderness, research has shown that visitor satisfaction is reduced to varying degrees depending on the types of encounters with other visitors, and that visitors report strong preferences for low levels of encounters [Stankey, 1973]. Therefore, managers of wildernesses receiving heavy use are beginning to take actions to modify or control use. Both the National Parks and the National Forests are rationing use of some areas. In some cases, managers have attempted to influence visitors to voluntarily shift their use to other areas or times through educational pamphlets and personal contacts.

However, all the managerial actions, except the establishment of rigid visitor itineraries, suffer from a major flaw — and rigid itineraries conflict with visitor freedom (which is also an important value) and are difficult to enforce. The manager's objective is to reduce use at overused locations and to avoid excessive levels of various types of encounters (on trails, at campsites, etc.), while reducing total use as little as possible. However, there has been no way to relate changes in total use or redistribution of use (either among entry points or to less-used times) to the number of encounters per party or to the amount of use of particular places within a wilderness. The complexity of travel routes, which characteristically overlap and intertwine, and the variability in travel decisions are so great that neither intuition nor analytic solutions are useful predictors of the variables of interest for a given amount of use.

Use pressures and encounters resulting from any given use level and pattern cannot be predicted and trial and error is not an effective approach to this problem because of the time required, risk of damage to resources, adverse public reactions, and so on. Therefore, a wilderness travel simulation model was developed by V. Kerry Smith of Resources for the Future and Webster and Heck of IBM to provide a better way to formulate and evaluate use management policies. [The model is coded in GPSS, and has been operated successfully on IBM 360 and 370 series, CDC 6600, Univac 1108 (Forest Service, Martin Wefald), and Boeing Computer System (National Park Service). Inquiries on use should be addressed to the first author.] Travel simulation models are common, but the requirements for the wilderness model were quite different. In particular, the interest in encounters was unique.

A GENERAL DESCRIPTION OF THE SIMULATOR

The computer program for the model generates different types of visiting parties (large, medium, and small groups of hikers and horseback riders in most applications) who arrive at the area at various simulated dates and clock times, enter at particular access points, select routes of travel, and move along them [Smith, Webster, and Heck, 1974; Smith and Krutilla, 1974]. The simulated parties may overtake and pass slower parties moving in the same direction (overtaking encounters), pass parties moving in the opposite direction (meeting encounters), or pass within sight of parties camped in areas visible from trails or other travel routes, such as rivers (visual encounters). Parties that stay overnight select campsites which they may share with other camping parties (camp encounters). Later, camping parties leave the campsite and continue on their chosen routes, and eventually leave the area.

To make the model operational, certain data are needed on the area and its recreational use. The travel network must be known, and basic facts about how different types of visitors behave within it — their patterns of arrival, various routes followed and relative popularity of each, travel speeds, and so on. This information is supplied to the model as probability distributions. The simulator provides both detailed and aggregated output information for each individual simulation of a particular use situation or "scenario."

The model has so far been applied in several land and river wildernesses. The present paper focuses on the application to the Desolation Wilderness in Calfornia, a high, mountainous, lake-dotted National Forest area of about 26,000 hectares that is very heavily visited. The use rationing plan begun in the Desolation Wilderness in 1978 was developed primarily with the aid of the simulation model. Daily entry quotas for each of 16 trailhead entry points were established on the basis of simulated use patterns that reduced encounters most efficiently; i.e., most reduced average encounters per party, per day, relative to the size of the reduction in total visitor numbers. Special sample surveys provided the needed input information on use and visitor behavior, as well as on encounter levels experienced, used for validity tests described below.

VALIDATION PROCEDURES

In a recent paper, Miller [1975] outlined a series of validation procedures which have often been used by investigators in the field of simulation. Miller's outline includes six validation tests:

1. User involvement
2. Content validity or "reasonableness of input" test
3. Face validity or "reasonableness of output" test
4. Sensitivity analysis
5. Historical goodness of fit test
6. Forecasting or predictive power test.

Each of these tests was used in the Desolation Wilderness application.

User Involvement and Content Validity

User involvement and content validity normally go hand-in-hand. The former is largely a necessary condition for the latter. User involvement means that a group of potential users — in this case, recreation managers — has been actively involved in the construction of the model from conceptualization to implementation. Content validity implies a high degree of realism built into the model via the postulated set of relationships among variables and parameters, as well as via the reliability of the data base.

Resource managers from the Forest Service, some responsible for managing wilderness recreation use in general and others directly responsible for the Desolation Wilderness, were intimately involved throughout the development of the simulator. Moreover, data collection for this first, large-scale application could not have been carried out without the cooperation and involvement of management in the field. This helped to guarantee that the visitor survey would yield a reliable data base.

Face Validity

Face validity concerns the output of the simulation model. Like the content validity test, it involves largely intuitive judgment on the part of both user and investigator. It is usually associated with queries of the sort: "Does the output of the model seem reasonable to you in light of your familiarity with the wilderness and the conditions which the model is supposed to represent?"

The simulator's output for the base-case run (representing typical midsummer, peak-season use) was carefully scrutinized. This run involved input information on use patterns derived from data collected during the 1974 use season. The model, in turn, was expected to reproduce the use and encounter patterns of the season. These results were evaluated by the recreation staff who judged them satisfactory compared with the real-world situation.

However, it is possible to attempt to improve upon this procedure by comparing the output of the model with known analytical results. This is a quantitative verification test [Naylor and Finger, 1967] which is designed to test the internal validity or logic of the model, as well as the reasonableness of its output. Hence, in a sense, it is a combination of the content and face validity tests.

Verification: Combining Content and Face Validity Tests

Let us assume a hypothetical trail network composed of only one trail segment (or a single campsite). It is then easy to calculate the maximum possible number of encounters between parties traversing the trail segment in any direction in a given time period (say, a day). For example, two parties, each traveling in a different direction, can have, at most, two encounters (one for each party). Three parties can have, at most, six encounters (each party could meet or overtake two other parties, making six altogether). Thus, the total number of encounters is equal to:

$$\frac{N!}{(N-2)!2!} = (N-1)\mathrm{N},$$

where N is number of parties, while the number of encounters per party per day (or, in other words, per party-day) is $N-1$.

The linear relationship between encounters per party-day and use levels should hold exactly for camp encounters since in that case all the parties at a given campsite meet all the other parties camping there on a given night. It should be approximate for trails, since only a portion of the total parties using a particular trail segment will be traversing it during a given time interval. Clearly, the longer the trail segment and/or the longer it takes to traverse it (traveling in the uphill direction, for example), the closer will be the actual relation to the theoretical linear one. Furthermore, the higher the use level the more closely linear this relation will be. In real situations it is clearly unreasonable to expect that every party meets all other parties present in the area. But we may expect the result to continue to hold in a proportional way, say, $E = \rho (N - 1)$, where E is the number of encounters per party-day, and ρ is a parameter. We do not know exactly what determines ρ, or what its actual numerical value is, but one may speculate that it is a function of the trail network and the route system.

The verification test was carried out in the following manner:

After running the 1974 base case, a series of four experiments were performed in which use levels were increased or decreased by certain percentages. Using the base case as the zero point, two experiments involved increasing the use levels by 25 and 50 %, respectively, and the other two consisted of decreasing use levels by the same percentages. Each experiment was replicated four times as determined by sample size considerations. The results of these experiements show that the theoretical relation also holds quite well for the output of the Desolation Wilderness application. One may infer that the model is internally valid, and, at least with respect to relative encounter levels, its output is reasonable.

Sensitivity Analysis

A principal reason — though not the only one — for conducting sensitivity analyses relates to the testing of the model's validity. The investigator wants to examine whether the model behaves in an expected, reasonable manner given changes in some key input variables. We would normally have some *a priori* notion concerning the direction of change in the model's output from experience and previous observations on the real system. In any large-scale simulation model, there would clearly be a rather large number of possible sensitivity tests. We have carried out two such tests for the simulator — one reported in the previous section, the effect of changes in use levels on a key output variable, the number of encounters per party-day — and a second centered around the variability of transit time, measured by its coefficient of variation (CV).

Transit time is a major stochastic input variable in the model, and therefore it is important to test whether changes in the CV affect encounters in a logical way. In scenarios with a lower CV, more parties traveled at speeds closer to the modal speed; hence, one would expect a reduction in both the mean number of encounters and its variance. Overtaking encounters would be reduced particularly by a reduction in transit-time variability. In the extreme case, where $CV = 0$, parties of the same size and type cannot overtake one another, although faster types could still overtake slower types. The results of the test strongly suggest that the model is quite sensitive to changes in transit time variability and the changes are logical in direction and magnitude. Also, as we should have expected, the statistical tests show that reduction of transit time variability has no significant effect on camp encounters. The results of the statistical test are detailed in Schechter and Lucas [1978].

Goodness-of-Fit Test

Writing on the subject of validating computer simulation models, Naylor and Finger [1967] stressed the need for criteria which would indicate whether or not the output of a simulation model agrees sufficiently with historical observations. They strongly recommend that "specific measures and techniques must be considered for testing the 'goodness of fit' of a simulation model" (p.97). We therefore tested the degree of correspondence between the simulated output and visitor trip diary records using the nonparametric Kolmogorov-Smirnov (K-S) test.

The trip diaries asked visitors to indicate the number of trail and camp encounters they had on each day of the trip. The pertinent question is whether the set of observations generated by the simulation model and those collected in the field could come from similar parent distributions.

Overall, trail encounters reported by visitors agreed with simulated results fairly well — better than do camp encounters — although the data did reveal some discrepancies between simulated and visitor-reported trail encounters. The K-S test (selected because the distribution of daily party-day trail or camp encounters could not be characterized as being normal, and also because it tests the entire frequency distributions) indicated that, at the .01 level, there was no significant difference between the distributions for trail encounters, while there was a significant difference for camp encounters.

Thus we cannot conclusively reject the hypothesis that the two samples — one simulated, the other reported by visitors — come from the same distribution. It should be noted that the application of the K-S test required that sample values substitute for the parameters of the fitted distribution. Without adjustment this could possibly undermine the validity of the conclusions in either direction [Fishman, 1973]. Moreover, the discrepency for camp encounters is consistent and easily corrected by refining input data. The major reason for this discrepancy was that the camp areas defined for the simulator were too large, and thus the simulator tallied camp encounters that visitors did not report, either because they were unaware of the other camper groups, or because they felt some of the groups were so far away that they could or should be ignored. It is also likely that some camper groups made a special effort to camp out of sight of others. The use of a relatively simplified trail system also could have inflated camp encounters compared to visitor experiences, if real-world campers used substantially more camp areas than were specified in the model.

It should be emphasized that such discrepancies do not invalidate *relative* comparisons among different simulation runs. Improvement or deterioration in visitors' encounter experiences can still be identified reliably.

Forecasting or Predictive Power Tests

The ultimate test of a model, especially one designed to aid decision makers in policy formulation and analysis, concerns the model's ability to predict the pattern of events which would result whenever a policy that is tested in the model is later put into effect in the real system. Such a test is operationally the most precise interpretation, in the scientific sense, of the term validation. It obviously cannot be conducted unless the implemented policy was previously simulated by the model and observations of simulated events are available for comparison with observations of corresponding events in the real system. Therefore, Emshoff and Sisson [1970, p. 205] conclude

> ... a model is completely valid only when it has been demonstrated to be a reliable and accurate predictor of event sequences and value variations. Furthermore, a model is *useful* only when the decision maker believes (rightly or wrongly) that it is valid in this sense.

It is too early a stage in the process of total evaluation of the model to conclude how the simulator fares on this test. A definitive answer will have to wait until the model has been applied more often and under varied circumstances and situations, although experience to date is encouraging. In addition to being used to develop a use rationing system for the Desolation Wilderness, the simulator has been applied successfully to the recreational use of the Green and Yampa Rivers in Dinosaur National Monument in Colorado [Lime, Anderson, and McCool, Chap. 9, pp. 153-153 in Shechter and Lucas, 1978] and to the Moose and Missinaibi Rivers in Ontario [Carls, 1978]. Professor Daniel Stynes, Michigan State University, has adapted the model to boating on lakes, with a concern for congestion as related to safety (personal communication), and Professor Robert Manning of the University of Vermont has applied the model to use of the Appalachian and Long Trails in Vermont [Manning and Ciali, 1979], as well as for teaching. It has also been used in teaching at several universities, including the University of Idaho and Washington State University. The model is being used now to help develop a use management plan for the Maroon Bells-Snowmass Wilderness in Colorado and for Yosemite National Park in California.†

Other applications of the simulation model are in early stages, gathering input data. These include the Mount Jefferson Wilderness, Oregon, the Colorado River in Grand Canyon National Park, and Bowron Lakes Provincial Park in British Columbia.

Conducting predictive power tests will not be easy, however. For example, actual use will never be exactly like the simulated pattern, but goodness-of-fit tests, such as those described in the previous section, might prove useful for this purpose, especially in comparing simulated and real encounter and use distributions.

The validity tests described in this paper seem to have contributed a substantial degree of confidence to the model. Armed with such tools, the investigator can approach the task of implementing the simulation model not as an act of faith, but as an integral part of a scientific investigation.

REFERENCES

Carls, E. G., 1978, "A simulation model of wild river use," *Leisure Sciences* Vol. 1, No. 3, pp. 209-218.

Emshoff, J.R. and Sisson, R.L., 1970, *Design and Use of Computer Simulation Models*, Macmillan, New York.

Fishman, G.S., 1973, *Concepts and Methods in Discrete Event Digital Simulation*, John Wiley, New York.

Hendee, J.C., Stankey, G.H., and Lucas, R.C., 1978, *Wilderness Management*, GPO, Washington DC.

Herman, C.F., 1967, "Validation problems in games and simulations with special reference to models of international politics," *Behavorial Science* Vol. 12, pp. 216–231.

Lloyd, R.D. and Fischer, V.L., 1972, *Dispersed Versus Concentrated Recreation as Forest Policy*, 7th World Forestry Congress, Buenos Aires.

†*Ed. Note:* I am happy to attest that verification letters on these public sector applications of the model have been received from: D.H. Anderson, E.G. Carls, D.J. Stynes, R.E. Manning, G. Haas, and T. Bell (for Maroon Bells-Snowmass Wilderness), and J.W. van Wagendonk (for Yosemite National Park).

Manning, Robert E. and Charles P. Ciali, 1979, "The computer hikes the Appalachian Trail," *Appalachia* Vol. 43, No. 1, pp. 75-85.

Miller, D.R., 1975, "Recent developments in approaches to validation of simulation models,"paper presented to the ORSA meeting at Las Vegas, Nevada, November.

Naylor, T.H. and Finger, J.M., 1967, "Verification of computer simulation models," *Management Science* Vol. 14, pp. 92–106.

Shechter, M., 1975, *Simulation Model of Wilderness-Area Use; Model-User's Manual and Program Documentation*, National Technical Information Service, Springfield, Virginia, Order No. PB 251635.

Shechter, M. and Lucas, R.C., 1978, *Simulation of Recreational Use for Park and Wilderness Management*, Johns Hopkins University Press, Resources for the Future, Baltimore.

Smith, V.K. and Krutilla, J.V., 1974, "A simulation model for management of low density recreational areas," *Journal of Environmental Economics and Management*, Vol. 1, pp. 178–201.

Smith, V.K., Webster, D., and Heck, N. 1974, "Analyzing the use of wilderness," *Simulation Today* Vol. 24, pp. 93-96.

Stankey, G.H., 1973, *Visitor Perception of Wilderness Recreation Carrying Capacity*, USDA Forest Service, Intermountain Forest and Range Experiment Station, Ogden, Utah, Research paper No. INT-142.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1980—777-028/63 Region No. 8

PURCHASED BY USDA FOREST SERVICE
FOR OFFICIAL USE ONLY

INTERFACES
Vol. 10, No. 5, October 1980
Copyright © 1980, The Institute of Management Sciences
0092-2102/80/1005/0011$01.25

VALIDATING A LARGE SCALE SIMULATION MODEL OF WILDERNESS RECREATIONAL TRAVEL*

Mordechai Shechter
Department of Economics, University of Haifa, Israel

and

Robert C. Lucas
USDA Forest Service, Forestry Sciences Laboratory, Drawer G, Missoula, Montana 59801

ABSTRACT. A large-scale simulation model of the use of outdoor recreation areas, especially ones with dispersed recreation patterns, has been developed that provides a means for experimenting with modifications of use or area conditions to determine effects on use patterns and congestion. The model and its applications to date, especially to the Desolation Wilderness in California, are briefly discussed.

The main purpose of this paper, however, is to illustrate a validation procedure for such a large-scale simulation model, employing a series of validity tests. Because the question of validation has always constituted a thorny problem in the field of simulation, the paper attempts to show — in the context of a case study — how tests can go beyond the customary "tests of reasonableness" often employed in such cases.

The question of the validity of simulation models is often raised and discussed in the literature on simulation. The problem is simply this: To what extent can a manager trust the results of a simulation model as a guide in solving problems in the real system?

Unfortunately, there is no universal agreement on the best method(s) to test validity of a model, whether analytic or simulation. Hermann [1967] has suggested five different tests for models whose predictive powers have not yet been tested. Only one truly qualifies as a scientific, objective test. It measures the degree of correspondence between the model's prediction of observable events and the events themselves when they occur (under a policy which was simulated beforehand). If a model cannot be shown to be highly reliable in terms of construction and predictive ability, and if the prospective users of the model cannot be convinced of its reliability as a management tool, then the researcher's efforts have been fruitless.

Our purpose in this paper is to illustrate an approach employing a series of reasonable tests — not merely a test of reasonableness — to check the validity of a

---

SIMULATION; RECREATION

*This research was carried out under a cooperative agreement between Resources for the Future, Inc. and the U.S. Forest Service, which provided funds for the project. The first author was a visiting scholar at Resources for the Future when he was associated with this project. Portions of this paper are based on the book, *Simulation of Recreational Use for Park and Wilderness Management*, by the authors, published by Johns Hopkins University Press for Resources for the Future.

large-scale simulation model which is now being used by public land managers. Our basic aim has been to demonstrate a method to overcome the problem of validation of such a model. We did not aim to present a transferable technique, necessarily; other models and situations might require another mix of strategies. We describe first the model itself and the real-life problems it is designed to solve.

WILDERNESS RECREATION MANAGEMENT PROBLEMS

The number of people visiting most kinds of outdoor recreation areas has been growing as long as records have been kept. This growth often creates problems for management, with the nature of the problems depending on the type of area and the management objectives established for it. In the United States, growth in use of dispersed recreation areas has been rapid [Lloyd and Fischer, 1972], and the resulting problems of congestion and resource damage have been difficult for managers to solve. This is especially true in wilderness areas which, by law, are supposed to contain substantially unmodified, natural ecosystems and provide opportunities for solitude.

Studies of the distribution of wilderness recreational use show very uneven use patterns; they are heavily concentrated on certain fairly small portions of each area, while larger portions receive little use. Similarly, summer weekends usually experience sharp peaks in use. Redistribution of use in time and space seems to offer considerable hope for reducing the adverse effects of heavy use [Hendee, Stankey, and Lucas, 1978].

In addition to the legal objectives for wilderness, research has shown that visitor satisfaction is reduced to varying degrees depending on the types of encounters with other visitors, and that visitors report strong preferences for low levels of encounters [Stankey, 1973]. Therefore, managers of wildernesses receiving heavy use are beginning to take actions to modify or control use. Both the National Parks and the National Forests are rationing use of some areas. In some cases, managers have attempted to influence visitors to voluntarily shift their use to other areas or times through educational pamphlets and personal contacts.

However, all the managerial actions, except the establishment of rigid visitor itineraries, suffer from a major flaw — and rigid itineraries conflict with visitor freedom (which is also an important value) and are difficult to enforce. The manager's objective is to reduce use at overused locations and to avoid excessive levels of various types of encounters (on trails, at campsites, etc.), while reducing total use as little as possible. However, there has been no way to relate changes in total use or redistribution of use (either among entry points or to less-used times) to the number of encounters per party or to the amount of use of particular places within a wilderness. The complexity of travel routes, which characteristically overlap and intertwine, and the variability in travel decisions are so great that neither intuition nor analytic solutions are useful predictors of the variables of interest for a given amount of use.

Use pressures and encounters resulting from any given use level and pattern cannot be predicted and trial and error is not an effective approach to this problem because of the time required, risk of damage to resources, adverse public reactions, and so on. Therefore, a wilderness travel simulation model was developed by V. Kerry Smith of Resources for the Future and Webster and Heck of IBM to provide a better way to formulate and evaluate use management policies. [The model is coded in GPSS, and has been operated successfully on IBM 360 and 370 series, CDC 6600, Univac 1108 (Forest Service, Martin Wefald), and Boeing Computer System (National Park Service). Inquiries on use should be addressed to the first author.] Travel simulation models are common, but the requirements for the wilderness model were quite different. In particular, the interest in encounters was unique.

A GENERAL DESCRIPTION OF THE SIMULATOR

The computer program for the model generates different types of visiting parties (large, medium, and small groups of hikers and horseback riders in most applications) who arrive at the area at various simulated dates and clock times, enter at particular access points, select routes of travel, and move along them [Smith, Webster, and Heck, 1974; Smith and Krutilla, 1974]. The simulated parties may overtake and pass slower parties moving in the same direction (overtaking encounters), pass parties moving in the opposite direction (meeting encounters), or pass within sight of parties camped in areas visible from trails or other travel routes, such as rivers (visual encounters). Parties that stay overnight select campsites which they may share with other camping parties (camp encounters). Later, camping parties leave the campsite and continue on their chosen routes, and eventually leave the area.

To make the model operational, certain data are needed on the area and its recreational use. The travel network must be known, and basic facts about how different types of visitors behave within it — their patterns of arrival, various routes followed and relative popularity of each, travel speeds, and so on. This information is supplied to the model as probability distributions. The simulator provides both detailed and aggregated output information for each individual simulation of a particular use situation or "scenario."

The model has so far been applied in several land and river wildernesses. The present paper focuses on the application to the Desolation Wilderness in Calfornia, a high, mountainous, lake-dotted National Forest area of about 26,000 hectares that is very heavily visited. The use rationing plan begun in the Desolation Wilderness in 1978 was developed primarily with the aid of the simulation model. Daily entry quotas for each of 16 trailhead entry points were established on the basis of simulated use patterns that reduced encounters most efficiently; i.e., most reduced average encounters per party, per day, relative to the size of the reduction in total visitor numbers. Special sample surveys provided the needed input information on use and visitor behavior, as well as on encounter levels experienced, used for validity tests described below.

VALIDATION PROCEDURES

In a recent paper, Miller [1975] outlined a series of validation procedures which have often been used by investigators in the field of simulation. Miller's outline includes six validation tests:

1. User involvement
2. Content validity or "reasonableness of input" test
3. Face validity or "reasonableness of output" test
4. Sensitivity analysis
5. Historical goodness of fit test
6. Forecasting or predictive power test.

Each of these tests was used in the Desolation Wilderness application.

User Involvement and Content Validity

User involvement and content validity normally go hand-in-hand. The former is largely a necessary condition for the latter. User involvement means that a group of potential users — in this case, recreation managers — has been actively involved in the construction of the model from conceptualization to implementation. Content validity implies a high degree of realism built into the model via the postulated set of relationships among variables and parameters, as well as via the reliability of the data base.

Resource managers from the Forest Service, some responsible for managing wilderness recreation use in general and others directly responsible for the Desolation Wilderness, were intimately involved throughout the development of the simulator. Moreover, data collection for this first, large-scale application could not have been carried out without the cooperation and involvement of management in the field. This helped to guarantee that the visitor survey would yield a reliable data base.

Face Validity

Face validity concerns the output of the simulation model. Like the content validity test, it involves largely intuitive judgment on the part of both user and investigator. It is usually associated with queries of the sort: "Does the output of the model seem reasonable to you in light of your familiarity with the wilderness and the conditions which the model is supposed to represent?"

The simulator's output for the base-case run (representing typical midsummer, peak-season use) was carefully scrutinized. This run involved input information on use patterns derived from data collected during the 1974 use season. The model, in turn, was expected to reproduce the use and encounter patterns of the season. These results were evaluated by the recreation staff who judged them satisfactory compared with the real-world situation.

However, it is possible to attempt to improve upon this procedure by comparing the output of the model with known analytical results. This is a quantitative verification test [Naylor and Finger, 1967] which is designed to test the internal validity or logic of the model, as well as the reasonableness of its output. Hence, in a sense, it is a combination of the content and face validity tests.

Verification: Combining Content and Face Validity Tests

Let us assume a hypothetical trail network composed of only one trail segment (or a single campsite). It is then easy to calculate the maximum possible number of encounters between parties traversing the trail segment in any direction in a given time period (say, a day). For example, two parties, each traveling in a different direction, can have, at most, two encounters (one for each party). Three parties can have, at most, six encounters (each party could meet or overtake two other parties, making six altogether). Thus, the total number of encounters is equal to:

$$\frac{N!}{(N-2)!2!} = (N-1)\mathrm{N},$$

where N is number of parties, while the number of encounters per party per day (or, in other words, per party-day) is $N-1$.

The linear relationship between encounters per party-day and use levels should hold exactly for camp encounters since in that case all the parties at a given campsite meet all the other parties camping there on a given night. It should be approximate for trails, since only a portion of the total parties using a particular trail segment will be traversing it during a given time interval. Clearly, the longer the trail segment and/or the longer it takes to traverse it (traveling in the uphill direction, for example), the closer will be the actual relation to the theoretical linear one. Furthermore, the higher the use level the more closely linear this relation will be. In real situations it is clearly unreasonable to expect that every party meets all other parties present in the area. But we may expect the result to continue to hold in a proportional way, say, $E = \rho (N - 1)$, where E is the number of encounters per party-day, and ρ is a parameter. We do not know exactly what determines ρ, or what its actual numerical value is, but one may speculate that it is a function of the trail network and the route system.

The verification test was carried out in the following manner:

After running the 1974 base case, a series of four experiments were performed in which use levels were increased or decreased by certain percentages. Using the base case as the zero point, two experiments involved increasing the use levels by 25 and 50 %, respectively, and the other two consisted of decreasing use levels by the same percentages. Each experiment was replicated four times as determined by sample size considerations. The results of these experiements show that the theoretical relation also holds quite well for the output of the Desolation Wilderness application. One may infer that the model is internally valid, and, at least with respect to relative encounter levels, its output is reasonable.

Sensitivity Analysis

A principal reason — though not the only one — for conducting sensitivity analyses relates to the testing of the model's validity. The investigator wants to examine whether the model behaves in an expected, reasonable manner given changes in some key input variables. We would normally have some *a priori* notion concerning the direction of change in the model's output from experience and previous observations on the real system. In any large-scale simulation model, there would clearly be a rather large number of possible sensitivity tests. We have carried out two such tests for the simulator — one reported in the previous section, the effect of changes in use levels on a key output variable, the number of encounters per party-day — and a second centered around the variability of transit time, measured by its coefficient of variation (CV).

Transit time is a major stochastic input variable in the model, and therefore it is important to test whether changes in the CV affect encounters in a logical way. In scenarios with a lower CV, more parties traveled at speeds closer to the modal speed; hence, one would expect a reduction in both the mean number of encounters and its variance. Overtaking encounters would be reduced particularly by a reduction in transit-time variability. In the extreme case, where CV = 0, parties of the same size and type cannot overtake one another, although faster types could still overtake slower types. The results of the test strongly suggest that the model is quite sensitive to changes in transit time variability and the changes are logical in direction and magnitude. Also, as we should have expected, the statistical tests show that reduction of transit time variability has no significant effect on camp encounters. The results of the statistical test are detailed in Schechter and Lucas [1978].

Goodness-of-Fit Test

Writing on the subject of validating computer simulation models, Naylor and Finger [1967] stressed the need for criteria which would indicate whether or not the output of a simulation model agrees sufficiently with historical observations. They strongly recommend that "specific measures and techniques must be considered for testing the 'goodness of fit' of a simulation model" (p.97). We therefore tested the degree of correspondence between the simulated output and visitor trip diary records using the nonparametric Kolmogorov-Smirnov (K-S) test.

The trip diaries asked visitors to indicate the number of trail and camp encounters they had on each day of the trip. The pertinent question is whether the set of observations generated by the simulation model and those collected in the field could come from similar parent distributions.

Overall, trail encounters reported by visitors agreed with simulated results fairly well — better than do camp encounters — although the data did reveal some discrepancies between simulated and visitor-reported trail encounters. The K-S test (selected because the distribution of daily party-day trail or camp encounters could not be characterized as being normal, and also because it tests the entire frequency distributions) indicated that, at the .01 level, there was no significant difference between the distributions for trail encounters, while there was a significant difference for camp encounters.

Thus we cannot conclusively reject the hypothesis that the two samples — one simulated, the other reported by visitors — come from the same distribution. It should be noted that the application of the K-S test required that sample values substitute for the parameters of the fitted distribution. Without adjustment this could possibly undermine the validity of the conclusions in either direction [Fishman, 1973]. Moreover, the discrepency for camp encounters is consistent and easily corrected by refining input data. The major reason for this discrepancy was that the camp areas defined for the simulator were too large, and thus the simulator tallied camp encounters that visitors did not report, either because they were unaware of the other camper groups, or because they felt some of the groups were so far away that they could or should be ignored. It is also likely that some camper groups made a special effort to camp out of sight of others. The use of a relatively simplified trail system also could have inflated camp encounters compared to visitor experiences, if real-world campers used substantially more camp areas than were specified in the model.

It should be emphasized that such discrepancies do not invalidate *relative* comparisons among different simulation runs. Improvement or deterioration in visitors' encounter experiences can still be identified reliably.

Forecasting or Predictive Power Tests

The ultimate test of a model, especially one designed to aid decision makers in policy formulation and analysis, concerns the model's ability to predict the pattern of events which would result whenever a policy that is tested in the model is later put into effect in the real system. Such a test is operationally the most precise interpretation, in the scientific sense, of the term validation. It obviously cannot be conducted unless the implemented policy was previously simulated by the model and observations of simulated events are available for comparison with observations of corresponding events in the real system. Therefore, Emshoff and Sisson [1970, p. 205] conclude

> ... a model is completely valid only when it has been demonstrated to be a reliable and accurate predictor of event sequences and value variations. Furthermore, a model is *useful* only when the decision maker believes (rightly or wrongly) that it is valid in this sense.

It is too early a stage in the process of total evaluation of the model to conclude how the simulator fares on this test. A definitive answer will have to wait until the model has been applied more often and under varied circumstances and situations, although experience to date is encouraging. In addition to being used to develop a use rationing system for the Desolation Wilderness, the simulator has been applied successfully to the recreational use of the Green and Yampa Rivers in Dinosaur National Monument in Colorado [Lime, Anderson, and McCool, Chap. 9, pp. 153-153 in Shechter and Lucas, 1978] and to the Moose and Missinaibi Rivers in Ontario [Carls, 1978]. Professor Daniel Stynes, Michigan State University, has adapted the model to boating on lakes, with a concern for congestion as related to safety (personal communication), and Professor Robert Manning of the University of Vermont has applied the model to use of the Appalachian and Long Trails in Vermont [Manning and Ciali, 1979], as well as for teaching. It has also been used in teaching at several universities, including the University of Idaho and Washington State University. The model is being used now to help develop a use management plan for the Maroon Bells-Snowmass Wilderness in Colorado and for Yosemite National Park in California.†

Other applications of the simulation model are in early stages, gathering input data. These include the Mount Jefferson Wilderness, Oregon, the Colorado River in Grand Canyon National Park, and Bowron Lakes Provincial Park in British Columbia.

Conducting predictive power tests will not be easy, however. For example, actual use will never be exactly like the simulated pattern, but goodness-of-fit tests, such as those described in the previous section, might prove useful for this purpose, especially in comparing simulated and real encounter and use distributions.

The validity tests described in this paper seem to have contributed a substantial degree of confidence to the model. Armed with such tools, the investigator can approach the task of implementing the simulation model not as an act of faith, but as an integral part of a scientific investigation.

References

Carls, E. G., 1978, "A simulation model of wild river use," *Leisure Sciences* Vol. 1, No. 3, pp. 209-218.

Emshoff, J.R. and Sisson, R.L., 1970, *Design and Use of Computer Simulation Models*, Macmillan, New York.

Fishman, G.S., 1973, *Concepts and Methods in Discrete Event Digital Simulation*, John Wiley, New York.

Hendee, J.C., Stankey, G.H., and Lucas, R.C., 1978, *Wilderness Management*, GPO, Washington DC.

Herman, C.F., 1967, "Validation problems in games and simulations with special reference to models of international politics," *Behavorial Science* Vol. 12, pp. 216–231.

Lloyd, R.D. and Fischer, V.L., 1972, *Dispersed Versus Concentrated Recreation as Forest Policy*, 7th World Forestry Congress, Buenos Aires.

†*Ed. Note:* I am happy to attest that verification letters on these public sector applications of the model have been received from: D.H. Anderson, E.G. Carls, D.J. Stynes, R.E. Manning, G. Haas, and T. Bell (for Maroon Bells-Snowmass Wilderness), and J.W. van Wagendonk (for Yosemite National Park).

Manning, Robert E. and Charles P. Ciali, 1979, "The computer hikes the Appalachian Trail," *Appalachia* Vol. 43, No. 1, pp. 75-85.

Miller, D.R., 1975, "Recent developments in approaches to validation of simulation models," paper presented to the ORSA meeting at Las Vegas, Nevada, November.

Naylor, T.H. and Finger, J.M., 1967, "Verification of computer simulation models," *Management Science* Vol. 14, pp. 92–106.

Shechter, M., 1975, *Simulation Model of Wilderness-Area Use; Model-User's Manual and Program Documentation*, National Technical Information Service, Springfield, Virginia, Order No. PB 251635.

Shechter, M. and Lucas, R.C., 1978, *Simulation of Recreational Use for Park and Wilderness Management*, Johns Hopkins University Press, Resources for the Future, Baltimore.

Smith, V.K. and Krutilla, J.V., 1974, "A simulation model for management of low density recreational areas," *Journal of Environmental Economics and Management*, Vol. 1, pp. 178–201.

Smith, V.K., Webster, D., and Heck, N. 1974, "Analyzing the use of wilderness," *Simulation Today* Vol. 24, pp. 93-96.

Stankey, G.H., 1973, *Visitor Perception of Wilderness Recreation Carrying Capacity*, USDA Forest Service, Intermountain Forest and Range Experiment Station, Ogden, Utah, Research paper No. INT-142.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1980—777-028/63 Region No. 8