

MIT MODEL ANALYSIS PROGRAM:
What We Have Learned About Policy Model Review

David O. Wood
Center for Energy Policy Research
Massachusetts Institute of Technology
Cambridge, MA 02139

1. INTRODUCTION

The availability of computers, together with important developments in econometrics and the social and policy sciences, has stimulated development of formal models to aid in formulating, analyzing, and implementing better public policy. Important examples may be cited in energy, the environment, education, transportation, and economic policy. While the investment of modeling resources and effort in these areas has been substantial, the contribution of formal models to better public policy has yet to be demonstrated in any significant way; certainly the contribution has not equalled the anticipation of the policy modelers themselves.

One reason for relatively modest accomplishment may be the limited scope for model application. Important public policy issues involve--almost by definition--controversies and conflicts of values and/or economic interest between politically important constituencies. Policy models can never resolve these conflicts; at best, they serve as tools of analysis to be used/accepted by the constituencies in conflict. Another contributing factor to the limited influence of formal policy models may be that not enough effort is devoted to establishing the credibility of these models, and providing evidence to the various policy constituencies that the models are relevant and can be "trusted."

A typical feature of policy analysis/support models is that while the scientific research upon which they are based is reported in the scholarly literature, the models themselves are usually documented in the technical report literature. Hence these models do not receive the peer review provided in the scientific literature. Yet, the more important the underlying policy issue, the more important it is to ensure that the models are "ventilated" and reviewed by competent scholars in the relevant disciplines. Models not subjected to this process will be handicapped in several important ways. First, they will not receive support from the technical community. Second, opportunities to identify and correct "treatable" deficiencies and issues requiring further research may not be recognized and acted on in a timely manner. Third, and perhaps most important, "un-refereed" models will complicate the model sponsor's task of establishing the model's credibility to policy analysts and non-technical policy constituencies. This is a critical task whose importance has not received sufficient attention in the policy

sciences.

The MIT Model Analysis Program (MIT/MAP) promotes model review as an important element in the policy research process, and as one means of increasing the credibility of formal policy models. The program was initiated in 1978 and has completed a number of model reviews and related research. Table 1 presents a chronology of significant model reviews and related activities.

Table 1: MIT Model Analysis
Program Studies

Model Evaluations

Baughman/Joskow Regionalized Electricity Model [9]
Decision Focus, Inc. Over/Under Capacity Planning Model [7]
ICF, Inc. Coal and Electric Utilities Model [6]
Load Management Strategy Testing Model, Binomial Approximation Module [12]
Teknekron Research, Inc. TELEPLAN Electric Utility Model [8]
Oak Ridge National Laboratory Residential Energy Use Model [11]
URGE Advanced Utility Simulation Model [1]

Modeling Studies and Data Base Evaluations

Mellon Institute's Least-Cost Energy Strategy Study [2]
EIA's National Coal Model in the 1978 Annual Report to Congress [17]
EIA's International Energy Evaluation System [10]
Comparative Evaluation of Manufacturing Factor Demand Studies [3,4,5,15]

Other

Recommendations to EIA on Model Documentation, Public Access, and Evaluation [16]
Surveys of Model Evaluation Methods and Studies [13,14]

The first MIT/MAP study reviewed two important energy policy models--the Baughman-Joskow Regionalized Electricity Model and the Wharton Long-Term Energy Model (Kuh-Wood [1979]). Chapter 1 of that report,

analyze[d] these assessment case study experiences to identify key organization-

al procedural issues that must be addressed in the assessment process, and to develop a deeper understanding of the approaches to, and objectives of, policy model assessment (Kuh-Wood [1979], p.S-2).

I first consider how experience since the earliest study has affected our views regarding approaches to policy model review, procedural issues, and the organization and conduct of model review studies. I then go on to provide a more impressionistic view of what has been learned from these studies.

2. APPROACHES TO MODEL REVIEW

An important contribution of the first MIT/MAP study was a classification (see Figure 1) of approaches to model review. This classification emphasized two dimensions including the depth of analysis (descriptive, analytical, experimental) and the materials employed in evaluation (published literature, model applications, computer codes, operating computing programs). Four approaches were defined including:

- literature review: essentially descriptive, using published model reports and materials, and information from the open literature;
- overview: more analytical than the literature review, and using, if appropriate, the model computer code, and model applications;
- independent audit: essentially experimental in that the reviewer has access to the operating model via an intermediary, and so can design, execute, and evaluate computational experiments to resolve issues raised by descriptive and/or analytical methods; and
- indepth: similar to the independent audit, except that the reviewer has full control of the operating model.

This classification has proven quite useful in organizing our review studies. In retrospect, I would amend it only slightly, and then mostly in matters of emphasis. First, I would now make more explicit that the level of review should correspond to the stage of model development. For example, literature reviews are most appropriate in the early stages of model development when the model formulation is under consideration. Overview and independent audit reviews are most useful for newly developed models that are likely to undergo further development. Such reviews will usually raise more questions than can be answered, but these will be questions which must be considered in the subsequent development process. Finally, in-depth assessments are most appropriate for mature models whose development process is essentially complete. Here the earlier questioning of analytical and descriptive reviews has been dealt with; now it is performance that matters.

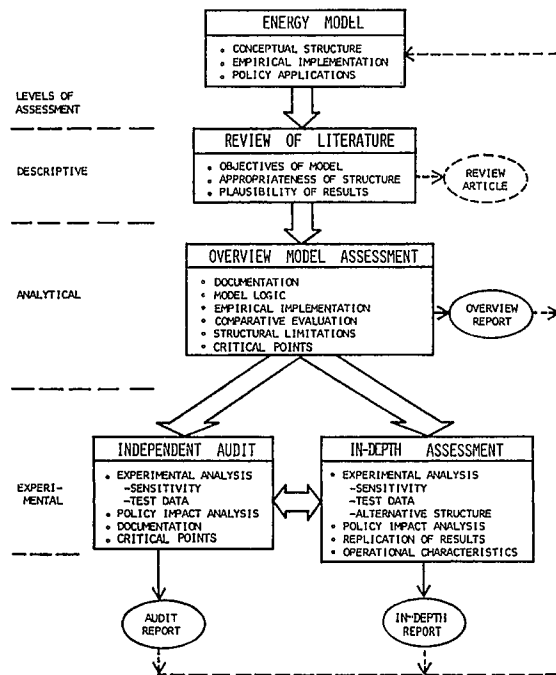


Figure 1: Approaches to Energy Model Assessment

Second, the focus of literature review on a specific model now seems too narrow. I would now amend this approach to emphasize formulating policy issues of interest, and evaluating the availability and status of scientific information and methods to support development of models to analyze these issues. This approach blurs into what we usually describe as the objective of policy model formulation. To my knowledge, very few reviews of this type have been published and discussed in advance of model development and research.

Third, the distinction between descriptive and analytical reviews now seems to contribute little of operation interest. The important distinction is descriptive and analytical versus experimental; this is what matters most, both in terms of the questions that can be posed and answered, and for the the organization and conduct of the review.

Finally, the need to approach reviews sequentially--an idea implicit in the earlier analysis--should be emphasized. As importantly, each stage of the review process should be formally completed via published reports and with modeler commentary before the next stage is undertaken.

3. PROCEDURAL ISSUES IN ORGANIZING AND CONDUCTING POLICY MODEL REVIEWS

In Kuh-Wood (1979), we proposed procedural guidelines for the organization and conduct of policy model reviews. These guidelines

were based primarily on our experience with Martin Baughman, co-developer with Paul Joskow of the Regionalized Electricity Model. These guidelines were summarized as follows:

- Assessor/modeler relations: A formal agreement should be reached defining the relationships between modeler and assessor with regard to,
 - resources to support modeler as well as assessor;
 - extent and nature of modeler/assessor interactions;
 - confidentiality of intermediate results;
 - opportunity for modeler response; and
 - post-assessment activities.
- Potential model applications: A wide-ranging list of potential applications of the model, incorporating suggestions from all interested parties, should be drawn up at an early stage to provide an explicit policy context for the assessment.
- Definition of a standard model: A standard version of the model must be agreed upon and "locked up" prior to the start of experimental analysis. It is desirable, however, to permit changes to be made during early stages of the assessment, particularly if the changes are to correct errors uncovered in the overview assessment.
- Assessors as modelers: Assessors can and should suggest ways in which the model can be improved, but they should not themselves implement the improvements. To do so would compromise the integrity of the assessment process and would put the assessors in competition with the modelers.

These guidelines--especially those concerning assessor/modeler relations--have served us well. Experience suggests, however, a number of refinements and changes in emphasis. First, the original guidelines do not mention model and/or review sponsors. Sponsors of the MIT/MAP included EPRI, EIA, and NSF. In most cases, no special problems arose beyond those normally addressed in the contractual arrangement between the reviewers and their sponsors. The only exception occurred when the review sponsor was also the model sponsor. In that case, a critical evaluation reflected on the competence of the sponsor's internal review process causing some tension between the sponsor and the reviewer group. Even then, the review was completed without undue interference. It seems, however, appropriate to extend the above guidelines to include:

- Assessor/assessment sponsor relations: The agreement between the review group and their sponsor(s) should be sufficient to ensure the independence of the review group, and should provide sponsors with the opportunity to include dissenting/clarifying comments

in the final report in a form similar to that provided the modelers.

Second, the importance of obtaining modeler support and involvement in the review process simply can not be over emphasized. Providing financial resources and the opportunity to include written comments and rebuttals as a final chapter in the review report provides a discipline that eliminates many misunderstandings, and provides (i) clarification of the issues that the model is intended to address, (ii) information on the context of the model development, and (iii) unpublished model related materials on formulation, implementation, and modeler conducted assessments and evaluations of sensitivity and performance.

Third, the guidelines concerning "assessors as modelers" now seems less relevant than originally thought. The original context for this was a concern that an indepth review put the model completely in the control of the reviewers, so that they would be free to make changes and then commercialize the results. Given our research interest, the opportunity and/or desire to do this has never arisen.

Upon reflection, however, and assuming all legalities are addressed, it now seems quite reasonable to me that a model review would be the first step in extending and further developing an existing policy model. A recent example is the review of the Teknekron Utility Simulation Model by the Universities Research Group on Energy as the first step in extending that model into the EPA Advanced Utility Simulation Model (Ref. 18), a model that was then reviewed by the MIT/MAP (Ref. 1).

In summary, then, the procedural guidelines developed in Kuh-Wood (1979) have served us well in the conduct of subsequent review studies. Of the "adjustments" I propose, only the need to clarify reviewer/review sponsor relations is of any operational consequence.

4. WHAT WE HAVE LEARNED: SOME IMPRESSIONS

The MIT Model Analysis Program has been underway for several years, and some considerable experience has been gained in conducting several reviews of policy models. Several general observations may be made based on these reviews and our experiences. First, perhaps the most important gains in policy model credibility come from the documentation of modeler self-assessments including sensitivity studies, and candid discussions of the strengths and weaknesses of a particular model, and its limits of application.

Second, obtaining the support and involvement of modelers has proved essential to effective reviews. The modelers provide the most efficient way to eliminate misunderstandings, and to provide materials and other information not included in the formal documentation. Establishing formal contracts with modelers has proven a very effective means of acquiring this input.

MIT Model Analysis Program

It should be noted, however, that the "down-side" to modeler involvement, in addition to financial costs, includes time costs for the overall study (meetings, etc.,) and --perhaps more importantly--sets up opportunities for "negotiated" reviews particularly regarding summary reports and abstracts. Thus while modeler involvement is extremely important, it is also important to insure that modelers do not "infiltrate" the review process, and damage the independent nature of the review.

Third, without exception models that have been published in the peer review literature, or at least whose applications have been published in this literature, tend to be much better documented, and much more easily reviewed. The discipline of providing the background materials to support publication in the literature, while no guarantee of model quality, is an important indicator of the seriousness and maturity of the modelers, and the credibility of the model development process.

Fourth, evaluation of "poorly" documented models should stop at the point when the review group must actually begin to augment the formal documentation.

Fifth, as emphasized above, the approaches to evaluation should be sequential. An effort oriented to an experimental approach almost of necessity requires the reviewers to first conduct descriptive and analytical reviews. The discipline of preparing and discussing reports at these levels is critical in identifying and formulating clearly the issues to be pursued at the experimental level. This is, of course, best done when the review process is organized to complement the model development process.

Sixth, our experience suggests that full in-depth evaluations are much less cost-effective than originally thought. Most of the gains from an in-depth review can be obtained by designing computational experiments to be executed by the model operators. The costs of fully assimilating model operations is very high, especially for a university based research group such as the MIT/MAP.

REFERENCES

1. Baughman, Martin L., Ernst R. Berndt, James A. Fay, Dan S. Golomb, Charles D. Kolstad, Fred C. Schweppe, and David O. Wood [1986], "A Review of the State-Level Advanced Utility Simulation Model," MIT Energy Laboratory Report (MIT-EL-86-001), February.
2. Berndt, Ernst R., Michael Manove, and David O. Wood [1983], A Review of the Energy Productivity Center's Least-Cost Energy Strategy Study, (EPRI EA-2753) Electric Power Research Institute, Palo Alto, January.
3. Berndt, Ernst R., Robert B. Hirsch, and David O. Wood [1981], "An Intercountry Translog Model of Factor Substitution Responses: Comment," MIT Energy Laboratory Working Paper (MIT-EL 81-039WP) June.
4. Berndt, Ernst R. and David O. Wood [1981], "Engineering and Econometric Interpretations of Energy-Capital Complementarity: Reply and Further Results," American Economic Review, Volume 71, December.
5. Berndt, Ernst R. and David O. Wood [1979], "Engineering and Econometric Interpretations of Energy-Capital Complementarity," American Economic Review, Volume 69, June.
6. Energy Model Analysis Program [1981], "The ICF, Inc. Coal and Electric Utilities Model: An Analysis and Evaluation," (7 volumes) MIT Energy Laboratory Report (MIT-EL-81-015) Submitted to Electric Power Institute, October.
7. Energy Model Analysis Program [1981], "The Over/Under Capacity Planning Model Review and Evaluation," (2 volumes), MIT Energy Laboratory Report (MIT-EL-81-045) Submitted to the Electric Power Research Institute, January.
8. Gruhl, James [1982], "Development Assessment of the TERA Teleplan Electric Utility Integrated Planning Model," Report to the Electric Power Research Institute, Palo Alto, January.
9. Kuh, Edwin and David O. Wood [1978], Independent Assessment of Energy Policy Models: Two Case Studies, (EPRI Report EA-1071) Electric Power Research Institute, Palo Alto, May.
10. Labys, Walter C. [1982], "Evaluation of the International Energy Evaluation System," MIT Energy Laboratory Report to the Energy Information Administration.
11. McFadden, Daniel [1982], An Evaluation of the ORNL Residential Energy Use Model, (EPRI Report EA-2442) Electric Power Research Institute, Palo Alto, June.
12. Schweppe Fred C., and Michael Caramanis, Martin Baughman, and James Gruhl [1981], "Overview Assessment of the Binomial Approximation Module of the DFI, Inc. Load Management Strategy Testing Model," MIT Energy Laboratory Report Submitted to the Electric Power Research Institute, Palo Alto, December.
13. Wood, David O. [1981], "Energy Model Evaluation and Analysis: Current Practice," in NBS Special Publication 616, US Government Printing Office, Washington, DC, October.
14. Wood, David O. [1980], "Model Assessment and the Policy Research Process: Current Practice and Future Promise," in NBS Special Publication 569, US Government Printing Office, Washington, DC, February.
15. Wood, David O. and Robert B. Hirsch [1981], "Reconciling Econometric Studies

- of Factor Demand: Data and Measurement Issues," MIT Energy Laboratory Working Paper (MIT-EL 81-011WP), April.
16. Wood, David O. and Martha J. Mason [1981], "Recommendations Concerning Energy Information Model Documentation, Public Access, and Evaluation," MIT Energy Laboratory Report (MIT-EL-81-016), April.
 17. Wood, David O. and Martha J. Mason [1981], "Analysis Quality Report on the EIA Annual Report to Congress 1978, volume III: Coal Supply," MIT Energy Laboratory Report (MIT-EL-81-017), April.
 18. Universities Research Group on Energy [1981], "A Review and Critique of the Teknekron Utility Simulation Model," University of Illinois, December.