Proceedings of the 1997 Winter Simulation Conference
ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson

COVALIDATION OF DISSIMILARLY STRUCTURED MODELS

Samuel A. Wright
Kenneth W. Bauer, Jr.

Department of Operational Sciences
Graduate School of Engineering
Air Force Institute of Technology
2950 P Street
Wright-Patterson AFB, Ohio 45433-7765, USA

ABSTRACT

A methodology is presented which allows comparison between models under different modeling paradigms. Consider the following situation: Two models have been constructed to study different aspects of the same system. One model simulates a fleet of aircraft moving a given combination of cargo and passengers from an onload point to an offload point. A second model is an optimization (linear programming) model, which for given cargo and passenger requirements, optimizes aircraft and route selection in order to minimize late- and non-deliveries. The optimization model represents a much more aggregated view of the airlift system than does the simulation. The two models do not have immediately comparable input or output structures, which complicates a comparison of the two models' outputs. A methodology is developed to structure this comparison.

Two models which compare favorably using this methodology are considered *covalid* models. We define the covalidity of models in the *narrow sense* as models which perform similarly under approximately the same input conditions. Structurally different models which are covalid (in the narrow sense) may hold the potential to be used in an iterative fashion to improve the input (and thus, the output) of one another. Ultimately it is hoped that we may, through such a series of innovations, effect a convergence to valid representations of the real-world situation. We define such a condition as covalidation in the *wide sense*. Further, if one of the models has been independently validated (in the traditional meaning), then we may effect a *validation by proxy* of the other model through this process.

1 INTRODUCTION

Mathematical models provide representations of real-world systems. These representations may take many forms: simulation, optimization, or regression, to name a few. Each requires a set of parameters and independent variables as input, employs some set of rules and relations, and returns a (set of) dependent variable(s) as output. It is likely that a real-world system has more than one model type which is used to characterize some unique aspect of the system. For instance, a system which consists of implementing a schedule of events of variable duration could be modeled as a simulation in order to determine the (distribution of) total time required to run the schedule. Such a model is descriptive in nature and could be used to forecast resource or personnel requirements. Alternately, an optimization model may be created in order to determine the shortest length of time in which the schedule may be implemented. This model optimizes some aspect of the system and could be used to study policy changes, such as improvement of the very schedule driving the simulation model. Either model (or both) may be a valid representation of the real-world system in question, and each should give results appropriate for its respective purpose.

In order for models to be useful, they must accurately model the real-world system. Model validation is defined as "substantiating that the model, within its domain of applicability, behaves with satisfactory accuracy consistent with the study objectives" (Balci 1994). There are a myriad of techniques available to assess the validity of simulation models (see Balci 1994, Sargent 1996b, or Sargent 1996a for examples), and many of these techniques may be applied to more general classes of models. These techniques include both objective and subjective tests of models, assessing the validity of model assumptions, model structure, behavior of model execution, and the model's output performance, among other tests. Depending on the study, different aspects of model validation may be of paramount importance. This effort concentrates on models' output performances.

There is a large body of literature relating to determining the output validity of a model for which real-world data can be obtained (see Law and Kelton 1991, p. 314-322 or Balci 1994 for examples and further references). The use of similarly structured models (specifically, simulations) has been proposed to assist in establishing model credibility (Diener, Hicks, and Long 1992). Often, the system being represented cannot be sampled in order to make comparisons. In such a case, a method of model output validation is not as obvious.

In this paper, we propose a method of *covalidation*, possibly resulting in a *validation by proxy*, in which dissimilarly structured models (e.g., optimization and simulation) representing the same real-world system for which output data are unattainable may be contrasted.

The rest of this paper is organized as follows. The next section provides a background to the specific application which generated the requirement for this research. The following section introduces the surrogate models used to test the developed methodology. The next three sections present methods for model covalidation. The first of these defines relevant terms. The second describes an iterative technique designed to exploit the crossflow of information between dissimilarly structured models. The third section presents a marginal analysis approach to model comparison. Following these sections, the methodology is demonstrated using surrogate test models. We conclude the paper with a summary and recommendations for further research.

2 MOTIVATION

The ultimate goal of this research is to contrast two large scale models. One model is a large scale discrete event simulation model which enjoys a relatively high level of acceptance. The other is a large scale linear programming model designed to optimize the same general system modeled by the simulation. The simulation model is used by the Air Mobility Command (AMC) and is known as the Mobility Analysis Support System (MASS). The Airlift Flow Module (AFM) is the simulation core of MASS. It simulates the movement of detailed cargo requirements through the airlift system based on the availability of aircraft, air routes, and air base infrastructure and resources.

The other model to be studied is a large scale optimization model developed jointly by the Naval Postgraduate School (NPS) and the RAND Corporation known as the NPS/RAND Mobility Optimizer (NRMO) (Morton, Rosenthal, and Weng 1996). NRMO models the strategic airlift system as a multi-period, multi-commodity network-based linear program (LP) with many side constraints. Use of the model is intended to provide insight into mobility problems such as fleet and infrastructure adequacy, and the identification of system bottlenecks.

NRMO's minimization objective function is the weighted sum of three sub-objectives. First, a large relative weight is attached to the non-delivery of cargo. Next in relative importance is minimizing the lateness of cargo deliveries. Finally, a third, relatively small weight is applied to minimizing the penalty for performing certain undesirable actions (such as deadheading crews). The weights applied to each sub-objective are subjective, but are generally ordered in that the weight for non-delivery is far greater than the weight for late delivery which, in turn, is far greater than the weight of the penalty.

The two models have many differences as well as similarities. The similarity to be capitalized on in this research is that both model strategic airlift and provide certain common outputs, such as the amount of cargo delivered. A major difference between the models is that NRMO optimizes the airlift schedule while MASS schedules flights based on the next availability of aircraft and a prioritized list of routes. Also, MASS models most event durations as random variables, while NRMO only models a mean value. A third major difference is that while MASS provides a detailed look at many aspects of the airlift system, NRMO represents a much more aggregated view of the airlift system.

The MASS simulation is currently in use by AMC Studies and Analysis Flight (AMCSAF), and though a formal validation has not occurred, its results are generally accepted as valid. One obstacle to a traditional, output-based validation of either model is that there is no way to collect real-world data from a strategic airlift system due to the infrequency of actual large-scale conflicts. Desire by AMCSAF to have some basis for the use of the NRMO optimization model has provided a motive to compare and contrast the two models.

3 TEST MODELS

As an investigation into the feasibility of the proposed methodology, very small scale test models are developed which are based on the MASS and NRMO models. First, a small scale simulation is constructed which models the movement of simple blocks of cargo with a fleet of identical aircraft from a single onload point to one of two offload points, after which aircraft recover to the onload point for further missions. The proportion of movement to each offload is user defined.

A small LP is also created which optimizes the amount of cargo that can be moved across the same airlift system as the simulation. Input into this model is an estimate of the efficiency of ramp space usage, which should account for the fact that ramp space may not be optimally scheduled in practice. This LP outputs, along with the total amount of cargo moved, the amount of cargo moved to each offload, thereby implying an optimal proportion of use for the two offload points.

Data used in these models are notional only and are not intended to resemble any actual airlift system data. Likewise, results obtained from these models are not intended to mimic those of the MASS or NRMO models, or those of any actual airlift scenario.

4 DEFINITIONS

First, a framework for covalidation is established. In general, the *covalidation* of two (or more) models of similar or dissimilar structure representing the same real-world system is the process of comparing the models, mindful of each individual model's domain of applicability, with the object of relative substantiation. *Covalidity* can be thought of as a matter of degree. *Covalidity in the narrow sense* is defined as a description of models which perform similarly under approximately the same input conditions. The extent to which models may be covalid in the narrow sense further depends upon the purpose of the models. For models with similar purposes (but perhaps different levels of detail or different modeling paradigms, such as two simulations of the same system with different levels of aggregation), the concept of narrow sense covalidity applies directly. For models with dissimilar purposes (such as the comparison between an optimization and a simulation), the definition may imply the ability to effect a meaningful cross-flow of information between the models.

In the case of dissimilar models, it is assumed that certain outputs from one model may contain information useful as input to the other model. If such a condition exists, it is possible the models may be used iteratively to enhance the performance of one another, resulting in a type of output convergence. Demonstrating such an iterative scheme is the secondary focus of this research.

Covalidity in the wide sense is defined as a description of models which not only are covalid in the narrow sense, but also which can be shown to be valid representations of the real world system. Further, if one of the models has been independently validated from the perspective of its intended use, models that are covalid relative to that model are considered *valid by proxy*.

5 INFORMATION CROSSFLOW

In dealing with dissimilarly structured models, the differences between the models must be carefully examined and exploited. Typically, dissimilarly structured models not only have different input (including both variable and parameter) sets, but could also have different levels of aggregation, as well as different capabilities in terms of modeling the actual system. In order to make a reasonable assessment of the models' covalidity, however, each of these differences must be examined.

Different levels of aggregation are common between structurally different models. The MASS simulation models the airlift system to a high level of fidelity compared to the NRMO optimization model. When comparing such models, the designed level of aggregation should be maintained for each model. In other words, the fidelity of models should not be compromised for the sake of "fair comparison." Optimistic optimization results could prove to be the result of unwarranted aggregation. For example, the infinite divisibility of units of cargo in the optimization could result in more cargo movement than is actually possible. By maintaining the appropriate levels of fidelity in each model, the covalidation process may also provide information on the appropriateness of such aggregation.

Whether to use the different modeling capabilities which exist between models should be carefully considered on a case by case basis before model comparison is performed. Generally, the "extra" capabilities of one model compared to the other model should be switched off, if possible. For example, since NRMO can effectively model aerial refueling aircraft operations while MASS cannot, the NRMO capability should be turned off during the comparison. If this is not possible, arrangement of the input variables or parameters should be such that the capability has no effect.

An exception to this general rule occurs when differences inherent in the modeling paradigms used allow one model to adequately model a system aspect while the other cannot. A simple example of this is that NRMO does not model the variability inherent in many airlift processes, relying instead on averages, while MASS models the random distributions of such variables. This difference in capabilities accounts for a fundamental difference between the two models, one for which comparisons in terms of covalidation are desired.

A difficult area is ensuring a rough parity in inputs between the two model types. Even when a particular input feeds both models, if the level of aggregation for this input is different between the two models, special care must be taken to ensure equitable representation.

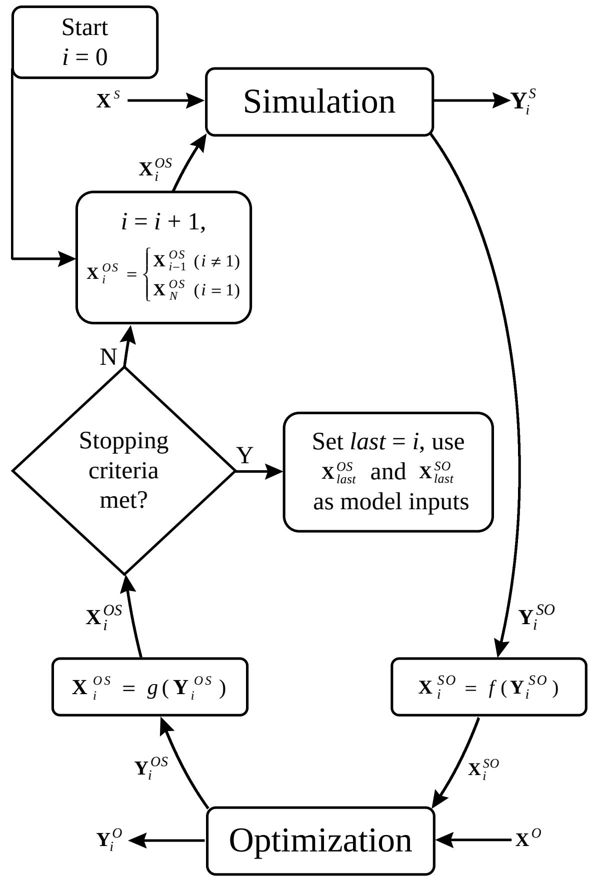
The real trick is to select outputs of one model which will serve as inputs to the other. In general, it is not clear that finding such information to crossflow is possible. However, since optimization models provide the "best solution" while simulation models can provide measures of system parameters, it seems reasonable that information could be meaningfully exchanged between these model types.

Here, an iterative method is employed in which inputs may converge. The purpose of this iterative scheme is to effect the crossflow of information between the models. That is, one model's output (or a function of that output) is supplied as input to the other model. For instance, NRMO provides as output the optimal selection of aircraft routes, while MASS accepts as input the frequency of route usage. On the other hand, MASS provides output which can be translated to the efficiency of parking space use, an input parameter required by NRMO. This iterative scheme is described in Figure 1. Though the figure specifically represents the application of simulation and optimization (e.g., MASS and NRMO, respectively) iteration, its concept is clear enough for general application.

In Figure 1, the i subscript denotes the current iteration. Each model has two input sets: $\mathbf{X}^S$ (or $\mathbf{X}^O$) is standard input for each model and $\mathbf{X}^{OS}$ (or $\mathbf{X}^{SO}$) is input which can be derived from or modified in reaction to the other model's output (input from the optimization to the simulation, or vice versa). For the first model execution, some (subjective) nominal values are placed on the inputs which, during later iterations, will be implied from output of the other model. Likewise, certain output, $\mathbf{Y}^S$ (or $\mathbf{Y}^O$), is not used by the other model while $\mathbf{Y}^{SO}$ (or $\mathbf{Y}^{OS}$) contains output which is utilized by the other model (output from the simulation to be used as input to the optimization or vice versa). The output subset that is used by the other model is "filtered" appropriately to make it usable as input for the other model. This filtering can be realized as a direct mathematical relationship or reflected as changes in policy or by adding model constraints.

Once every cycle, a check is made to determine if the stopping criterion has been met. This criterion can be that a model's input has converged. Alternatively, iterations may indicate that a point of diminishing returns has been reached with this process. Either way, the *last* iteration inputs are deemed those which are as close as possible, and they are used as the experimental design center for the ensuing model comparison.

Prior to developing the particulars of model comparison, a summary of the steps taken to achieve a crossflow of information is offered. First, the input/output structures of the models are carefully studied and appropriate adjustments are made due to the differences in level of detail and capability. Next,



In the figure:

i is the iteration number
last is the final iteration number
N implies a nominal value

Simulation input/output
$\mathbf{X}^S$ is input specific to simulation
$\mathbf{X}^{OS} = g(\mathbf{Y}^{OS})$ is input to simulation which is derived from output of optimization
$\mathbf{Y}^S$ is output of simulation
$\mathbf{Y}^{SO}$ is output from simulation which is to be used as input to optimization

Optimization input/output
$\mathbf{X}^O$ is input specific to optimization
$\mathbf{X}^{SO} = f(\mathbf{Y}^{SO})$ is input to optimization which is derived from output of simulation
$\mathbf{Y}^O$ is output of optimization
$\mathbf{Y}^{OS}$ is output from optimization which is to be used as input to simulation

Figure 1: Iterative Scheme

input/output links are determined which will allow for a potentially meaningful crossflow of information. Finally, the information provided by these links is used in an iterative fashion aimed at improving both models. Convergence of these inputs should indicate a state where the models, as closely as possible, represent the same scenario. Further, convergence in the model's outputs provides the first indication that the models are performing similarly. This is further examined in the following section.

6 MARGINAL ANALYSIS

In order to determine the level of model correspondence, a method of output comparison is required. Metamodels constructed across a relatively small experimental design of interesting and/or relevant input variables may provide a convenient means of effecting this comparison. Both models were exercised (during the iterative scheme) at the same basic input settings. This setting becomes the center point of an experimental design aimed at estimating the local gradients of each model. A comparison of these gradient estimates indicates whether the models respond in a similar fashion to perturbations in the selected inputs. However, complete agreement of these estimates between metamodels is not necessary for asserting that the models compare favorably.

A method is developed which allows comparison between the relative closeness of the model outputs as well as between estimates of gradients of the models representing the sensitivity of a selected output to a set of common inputs. Using this method, we are able to investigate the relative predictive value of the metamodels (through the output comparison) as well as compare the metamodels' abilities to provide a description of the physical system (through comparison of local gradients).

The next required task for model comparison is the creation of an experimental design. The design for this study is based on the desire to evaluate the sensitivities of a single output to changes in key inputs. The choice of design is dependent on the study. The number of model executions to be performed is restricted by the size of the models, the number of inputs to be varied, the desired design resolution, and the number of replications to be made at each design point (to reduce the output variability of simulations) (Box and Draper 1987).

The result of the iterative scheme is two sets of model inputs (to include parameters and variables), one for each model, which correspond nearly as possible to the inputs and outputs (as applicable) of the opposing model. The resulting convergence of certain inputs is valid only for the design point upon which the iterative scheme was performed. Assuming that the resulting inputs (whether variable, parameter, or logical rule) are constant across any large experimental design is incorrect in general. It is clear that at each design point across the common experimental design the iterative method may yield convergence to different sets of specific model inputs, and so this iterative method would require repeating at each point in a wider experimental design. For instance, a large change in the number of available aircraft (or any resource) may affect optimal routing. For this reason, we assume that design points are very close to one another and that the difference is unimportant.

7 ILLUSTRATIVE EXAMPLE

In order to demonstrate the use of the proposed methodology, an example is given using very small scale test models which represent the MASS and NRMO models. Data used in these small scale surrogates is notional only and no inferences should be made from the data or results to either the MASS and NRMO models or to any actual airlift scenario.

The scenario posed is that 50 similar aircraft must fly as many missions as possible to either of two airbases (A and B) in 15 days. The airbases can handle 5 and 10 aircraft maximum on the ground (MOG) at a time, respectively, and are different distances from the home base. Sensitivities of both the number of aircraft and the amount of available MOG are of interest here. In the simulation (Baby MASS) the flight times to (and from) each base and the ground time of the aircraft at bases A and B are random variables, while the optimization (Baby NRMO) assumes a mean value. (See Appendix for a more detailed account of both the simulation and optimization models.)

The simulation accepts as input a ratio for the use of the two bases, while the optimization yields optimal values for this ratio. Similarly, since it is not possible in practice to perfectly schedule the ground spaces available at the bases (as an optimization model would), the optimization accepts as input a MOG efficiency factor. The simulation may yield a practical maximum number of aircraft that can be serviced at a base from which a better estimate for MOG efficiency may be derived. The iterative scheme uses this information to attempt convergence to some "optimal" ratio of base visitation and MOG efficiency.

Table 1 shows the results of the iterative scheme with 10 runs being made at each iteration (for the simulation). A fifty-fifty split was used as the nominal ratio of aircraft sent to each base, and 1.0 was used as the starting MOG efficiency. When the simulation results indicated a bottleneck at a base, a MOG efficiency was calculated based on the number of aircraft that were actually able to be serviced at the base and the amount of time the aircraft had to be serviced at the base (see Appendix). For a Baby NRMO run, the number of planes routed to each base are determined. This ratio is used as direct input for the subsequent Baby MASS run (see Appendix).

Table 1: Iterative Scheme for Test Models

	Baby MASS			Baby NRMO		
i	A	B	Total	A	B	Total
	0.5	0.5	← %	eff. →	1.0	
1	84	91.4	175.4	52.0	140.0	192.0
	0.286	0.714	← %	eff. →	0.938	
2	50.7	129	179.7	56.8	131.3	188.2
	0.318	0.682	← %	eff. →	0.926	
3	57.1	127.3	184.4	57.8	129.6	187.4
	0.324	0.676	← %	eff. →	0.898	
4	57.3	123.4	180.7	60.0	125.6	185.6
	0.340	0.660	← %	eff. →	0.876	
5	59.7	120.4	180.1	61.7	122.6	184.3
	0.352	0.648	← %	eff. →	0.888	
6	62.6	122.1	184.7	60.7	124.3	185.1
	0.345	0.655	← %	eff. →	0.902	
7	62.2	124	186.2	59.7	126.2	185.9
	0.337	0.663	← %	eff. →	0.901	
8	60.3	123.9	184.2	59.7	126.2	185.9
	0.338	0.662	← %	eff. →	0.901	
9	60.3	123.9	184.2	59.7	126.2	185.9

As seen in Table 1, a stopping criterion is met since the Baby NRMO results converge in the 9th iteration. The input convergence is shown graphically in Figure 2. The final output values indicate that an average of 184.2 missions are flown in Baby MASS (with standard error of 2.21) and 185.9 missions are flown in Baby NRMO, and the comparison of these output values across the iterations is shown in Figure 3. The converged flying ratios and MOG efficiency are then used during the experimental design, for gradient estimation.

The metamodels were created using the number of aircraft and the amount of MOG at base B as independent variables perturbed over a 2^2 plus center point experimental design. (MOG at base B was selected since base B proved to be a system bottleneck.) The number of planes was varied plus and minus 10 percent (to 55 and 45 planes, respectively) and the MOG at base B was varied by plus and minus 1 unit of MOG (to 6 and 4 MOG units, respectively). The number of missions flown is the dependent variable.

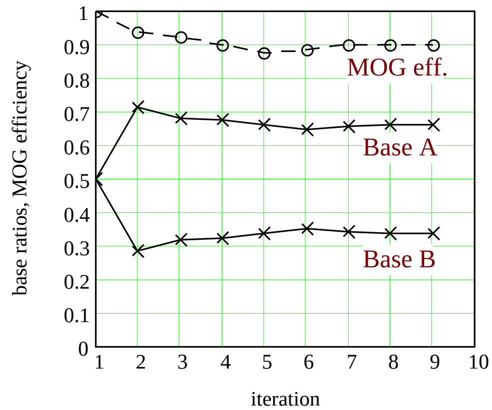


Figure 2: Base Ratio, MOG Efficiency Convergence

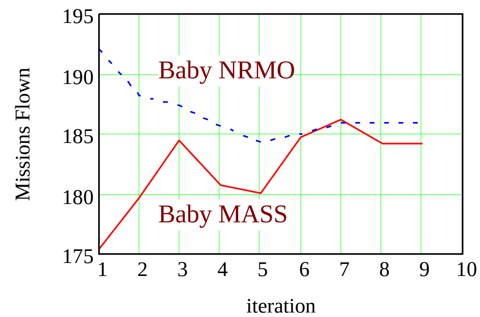


Figure 3: Thruput Improvement

The regression results are summarized in Table 2. Note that for the Baby NRMO model, the only error possible would come from specification bias. Since there is no error in Baby NRMO's metamodel, it is clear that none of the design points are outside the critical region found at the design's center, i.e., the optimal basis did not change at any design point. The metamodels are also shown graphically in Figure 4 and 5.

Table 2: Comparison of Test Models

	Baby MASS	Baby NRMO	Difference
Mean	177.5	185.9	8.4
Planes	12.7	13.0	.3
MOG at B	13.7	11.2	-2.5
Interaction	3.6	0	-3.6
coeff. std. err.	1.81	0	1.81
SSR	14418	11755	766
SSE	4159	0	4158
SST	18576	11755	4924
F*	53.2	∞	2.8
R^2	0.776	1	0.155

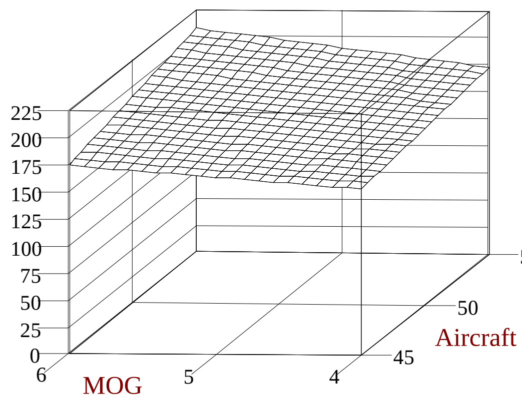


Figure 4: Baby MASS Metamodel

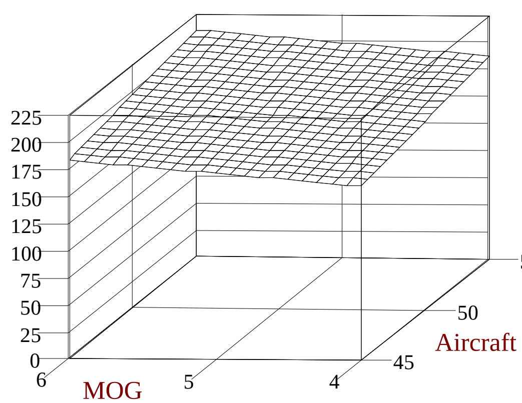


Figure 5: Baby NRMO Metamodel

The third column in Table 2 is formed by performing a regression of the differences between the Baby NRMO and the Baby MASS data. The relative insignificance of the overall comparison model (marginally significant at $\alpha = 0.05$) implies the similarity between the two models, with most of the difference reflected in the interaction term. However, the relatively large difference in mean values between the two models (compared to the center point arrived at through the iterative scheme) is a result of the Baby MASS model being evaluated across a relatively wide experimental region. The appropriate ratios of base A to B could be re-iterated at each design point or the design space could be narrowed in order to minimize this difference. Obviously, however, this is impractical in the case of the MOG at B factor, since the range is plus and minus one MOG unit, which cannot be fractionalized in the simulation.

8 CONCLUSIONS AND FUTURE RESEARCH

We have demonstrated a method which forms comparisons between models under different modeling paradigms. We have also introduced the notion of covalidity as two models which compare favorably using this methodology.

An insight taken from this research is the importance of not assuming that model starting conditions are static across wide experimental regions. In order to take advantage of our method of marginal analysis, keeping a narrow experimental region is vital since the input conditions arrived at in the iterative scheme likely have a narrow band of usefulness.

Further research is being conducted in this area, specifically the application of these covalidation methods to the actual MASS and NRMO models.

APPENDIX A

Figure 6 shows the basic network used in the test models. In the network for the base case (center of experimental design), 50 aircraft are sent from Home to either A or B. At A or B, the aircraft are unloaded and serviced, and they return to Home. Scenarios for both the simulation and the optimization cover 15 days. The flight and service time distributions shown are used in the simulation, but the optimization formulation only reflects the mean times.

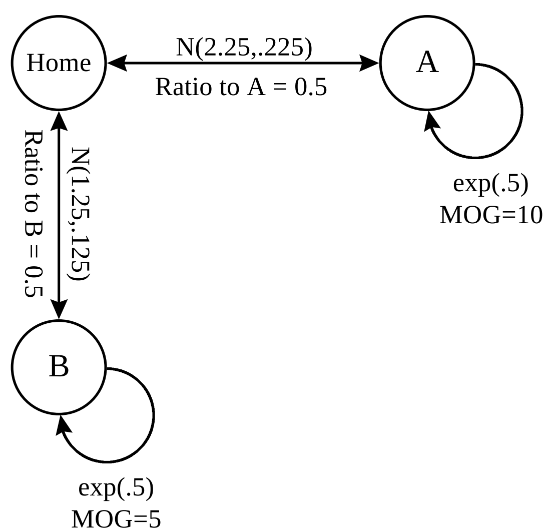


Figure 6: Test Model Network

The optimization formulation is shown below. The basic formulation maximizes the number of missions flown. Constraints are added which limit the number of planes utilized each day to 50 and the number of planes serviced by a base to some fraction (MOG efficiency) of the available MOG for the base. Not shown is an additional constraint which accounts for the amount of time required for the initial aircraft to reach the bases.

$$\text{maximize } thruput = \sum_t \sum_b \mathbf{X}(b,t)$$

subject to:

$$\sum_b \sum_{flying(t)} \mathbf{X}(b,\ flying(t)) \le plane \forall t$$

$$\mathbf{X}(b,t) \le MOGeff \times MOG(b) \forall b,t$$

$$\mathbf{X}(b,t) \ge 0$$

where

b is base A or B
t is day 1 through 15
X is number of missions arriving at *b* during *t*
flying is a vector which accounts for the flight days during a mission to *b* arriving at *t*
plane is total number of aircraft available
MOG is the maximum number of aircraft simultaneously serviceable by *b*
MOGeff is a measure of how efficiently *MOG* can be used if *b* is bottlenecked

The equations which derive the input of one test model from the output of the other are specified below. The equation which filters Baby MASS output into Baby NRMO input calculates the fraction of MOG which the simulation could actually use at a base, given that the base was bottlenecked throughout the simulation.

$$\text{MOG eff.} = \frac{\#\text{ missions}}{\#\text{ days}} \div \text{avail. MOG / day}$$
$$= \text{MOGeff missions / MOG}$$

The equation which filters Baby NRMO output to Baby MASS input determines the ratio at which the Baby NRMO optimization flies missions to each base (A and B). As a function of total thruput to a base over the entire fifteen day scenario, the number of missions flown per day is determined by how many days to which each base is flown. In this scenario, base A was flown to on 13 of the days, and base B was flown to on 14 days.

$$\frac{\text{A}}{\text{B}} = \frac{\text{Total to A}/\#\text{ days flown to A}}{\text{Total to B}/\#\text{ days flown to B}}$$

REFERENCES

Balci, O. 1994. Validation, Verification, and Testing Techniques throughout the Life Cycle of a Simulation Study, *Annals of Operations Research*, 53, pp. 121-173.

Box, G.E.P. and N.R. Draper. 1987. *Empirical Model-Building and Response Surfaces*, John Wiley & Sons, New York.

Diener, D.A., H.R. Hicks, and L.L. Long. 1992. Comparison of Models: Ex Post Facto Validation/Acceptance, *Proceedings of the 1992 Winter Simulation Conference*, Arlington, Virginia, pp. 1095-1103.

Law, A.M. and W.D. Kelton. 1991. *Simulation Modeling & Analysis*, 2nd edition, McGraw-Hill, New York.

Morton, D.P., Rosenthal, R.E., and Weng, L.T. 1996. Optimization Modeling for Airlift Mobility, *Military Operations Research*, 1, 4, pp. 49-67.

Sargent, R.G. 1996a. Some Subjective Validation Methods Using Graphical Displays of Data, *Proceedings of the 1996 Winter Simulation Conference*, Coronado, California, pp. 345-351.

Sargent, R.G. 1996b. Verifying and Validating Simulation Models, *Proceedings of the 1996 Winter Simulation Conference*, Coronado, California, pp. 55-64.

AUTHOR BIOGRAPHIES

SAMUEL A. WRIGHT is a Ph.D. candidate in the Department of Operational Sciences at the Air Force Institute of Technology. He received a B.S. degree in Astronautical Engineering from the United States Air Force Academy in 1989, and he received an M.S. degree in Operations Research from the Air Force Institute of Technology in 1995. His research interests include applied statistics and simulation modeling. He is a member of Tau Beta Pi.

KENNETH W. BAUER, JR. is a Professor in the Department of Operational Sciences at the Air Force Institute of Technology. His research interests include the statistical aspects of neural networks and simulation modeling.