

## A COMPARISON OF SERIAL AND PARALLEL PROCESSING SIMULATION MODELS OF PILOT WORKLOAD

Thomas F. Schuppe  
Department of Operational Sciences  
Air Force Institute of Technology  
Wright-Patterson AFB, OH 45433, U.S.A.

### ABSTRACT

This paper discusses and evaluates several options for modeling a pilot while he performs normal cockpit duties. Using a queueing analogy, with the pilot modeled as the server and the pilot's tasks as customers, two discrete-event simulation models were developed. The main issue examined was the relative accuracy of a model which processes tasks in series versus one which processes tasks in parallel. Each of these models employed a priority system with preemption. The serial model processed one task at a time while the number of tasks that could be processed simultaneously in the parallel model was dependent on the mix of tasks currently demanding the pilot's resources. Three different measures from the simulation models were evaluated as workload surrogates: pilot idle time, task interruption rate, and the simultaneous task rate. Workload model predictions were validated by comparison to workload measurements taken under identical conditions in a high-fidelity flight simulator. The models' validation process suggests that the technique has merit and continued investigation is warranted.

### 1. INTRODUCTION

At the present time, several methods exist to measure pilot workload. These include the Subjective Workload Assessment Technique (Reid, et al) and the NASA Bi-Polar method (Battiste and Hart). However, all of these methods require that a pilot's workload be measured while completing normal pilot tasks in an aircraft, flight simulator, or some other test apparatus. These measurements are acceptable for operational aircraft systems, but inadequate for systems under design. It is especially important to be able to estimate or predict a pilot's workload in a proposed aircraft cockpit so that problems can be eliminated before expensive hardware has been purchased. Therefore, a method which reliably predicts workload without the use of hardware would be very beneficial. This paper describes a discrete event simulation model used in a methodology to predict pilot workload.

The method described in this research uses the mathematical structure of the Subjective Workload Assessment Technique (SWAT) as a framework for making predictions. SWAT assumes workload has three dimensions: time load, mental effort load, and psychological stress load. According to Potter and Acton

(5), time load refers to the degree of task overlap or interruption. Mental effort load refers to the amount of attention or concentration required for task performance. Finally, psychological stress load refers to the combined emotional and physical factors, such as anxiety or confusion, which can otherwise affect subjective load.

Normally, SWAT measures workload after a flight by asking a pilot to subjectively assess workload in each of these three dimensions. To predict workload, the technique is modified so that mental effort load and psychological stress load are subjectively estimated without actually flying in a aircraft. While pilots have been able to assess the impact of mental and psychological factors on workload, Battiste and Hart have found that humans are inaccurate in estimating the impact of complex or unfamiliar timing elements. This inaccuracy suggests a need for some other method to assess the impact of task timing and rate changes on workload. In this paper a model is described to estimate workload in the time load dimension by using measures derived from a discrete-event simulation of a particular scenario. This estimate of workload in the time load dimension is then combined, using SWAT methodology, with the two subjective workload estimates in the mental effort load and psychological stress load dimensions to form a workload prediction. This paper focuses on the development and validation of the discrete-event simulation models used in this study. A more complete description of this entire effort can be found in Schuppe.

### 2. THE MODELS

Some similarities exist between the serial and the parallel processing models. In each model, the basic structure of a queueing system was used. The pilot is the server and the tasks he accomplishes are his customers. Customers are served according to a priority system. Higher priority tasks preempt lower priority tasks. Within a priority class, a first-come, first-served discipline is employed. Three task priority classes are used: operating tasks, monitoring tasks, and planning tasks. Operating tasks are those required for aircraft control and subsystem operation. This class has the highest priority and can tolerate very little delay. Monitoring tasks are those which require the pilot to scan inside and outside the cockpit to update his knowledge of the aircraft's status and situation. This second class can accept short delays without serious consequences. The lowest priority class, planning tasks, can be deferred into

idle periods for completion. Some tasks, such as aircraft control, are recurrent and are modeled as repeat customers. Other tasks, such as passing a navigation turn point, occur only at specified points in a flight and are modeled as non-repeating customers.

Both simulation models were written in SIMSCRIPT II.5 and both were run on an IBM 3081 and a VAX 11/785. Run times for 25 replications on the IBM were approximately 15 CPU seconds for the serial processing model and 21 CPU seconds for the parallel processing model. On the VAX, the times were 14 CPU seconds and 41 CPU seconds, respectively. Each model contains twelve separate event routines which create tasks to be accomplished by the pilot. Of these twelve event routines, seven are repetitive while the remainder occur only once. The inter-arrival time for each repetitive task is random and is determined by Monte Carlo sampling from a designated probability distribution. The time required for the pilot to complete each task is also random and is determined by Monte Carlo sampling from another appropriate probability distribution.

### 2.1 The Serial Model

This model assumes a pilot processes all tasks sequentially. That is, he can process only one task at a time and cannot begin a second task until the first is completed or preempted. Any task which arrives while the pilot is working on another task must have a higher priority for service to preempt the working task. If unable to preempt, the newly arrived task must wait until the old task is completed and the pilot is free to begin service. All tasks waiting for service are ranked by priority. Once a task has been preempted, it returns to the waiting line to await another chance for service. Preempted tasks do not use their original service time for subsequent service, but resample to determine a new service time. Resampling is done to reflect the changed aircraft state when the task re-enters service. It is assumed that previously accomplished portions of tasks have to be repeated, changed, or corrected. Figures 1 and 2 display the task arrival and service completion processes respectively for the serial model. The primary estimators of Time Load produced by the serial model are pilot idle time and the ratio of interrupted tasks to task arrivals. These two estimators are used as surrogates, or replacements, for workload in the Time Load dimension of the SWAT methodology.

### 2.2 The Parallel Model

Conceptually, the serial model is straight-forward and easy to understand and implement. However, the parallel processing model is not as simple. It is not clear how many parallel tasks can be handled simultaneously by a human, or how performance on each task might be affected by the inclusion of additional parallel tasks. It is not too difficult to imagine situations in which a human may, or may not, efficiently accomplish two distinctly different tasks simultaneously. However, it is much more difficult to model this process and recognize task combinations

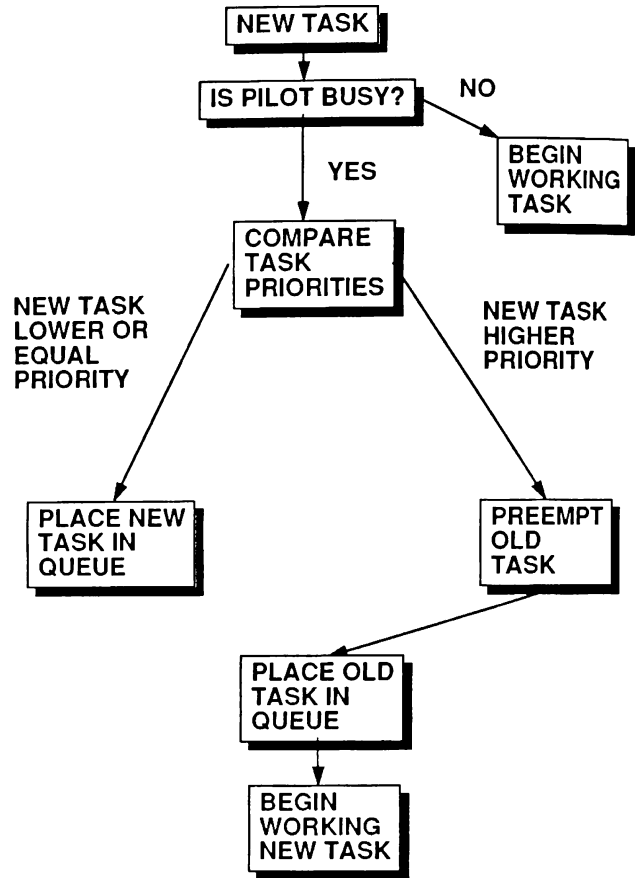
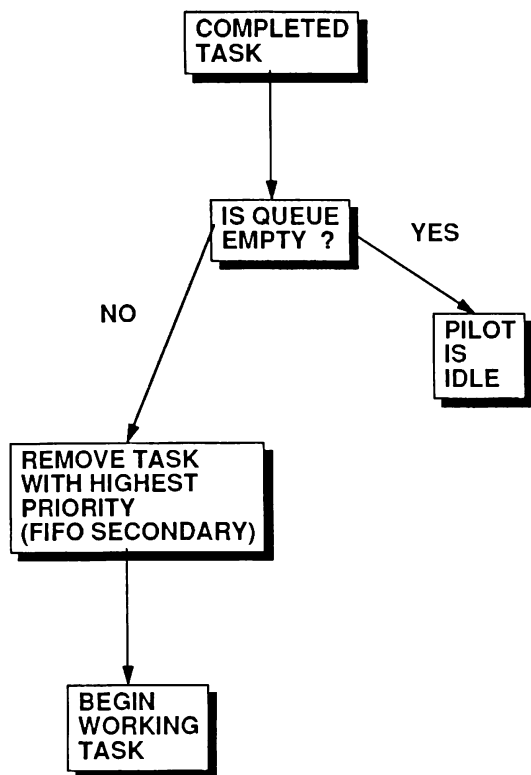


Figure 1: Task Arrival Process, Single Server Model

that can be done simultaneously, and others which cannot.

For those tasks accomplished simultaneously, degradation in terms of increased service times must also be recognized. For example, consider the situation of a human driving a car on a busy expressway. It is easy to imagine someone doing this task very competently while simultaneously accomplishing a second task, such as conducting a conversation with someone else in the car. This particular combination of tasks can be accomplished simultaneously because the tasks demand different resources from the driver: visual input and control output in one case vs. aural input and verbal output in the other. A small performance degradation may result from parallel completion of these tasks because both require some mental processing. This situation can be compared to another where the driver is attempting to read a newspaper while driving in heavy traffic. Both tasks require concentrated visual input and serious performance degradation will occur as the driver attempts to switch between the two tasks.

To represent this concept, a model was constructed which accomplished parallel processing of pilot tasks.



**Figure 2: Service Completion Process, Single Server Model**

The number of parallel channels available was not fixed, but was dependent on the pilot resources required by the tasks demanding service. Wickens' multiple resource model is used to define available pilot resources. This model postulates that humans possess several different capacities with resource properties. Task interference will occur if resources are shared, but little interference will occur if different tasks require different resources. Wickens' model describes three resource dimensions with binary properties: INPUT MODALITY (visual or auditory), PROCESSING STAGES (encoding and central processing or responding), and PROCESSING CODES (spatial or verbal). This results in eight (2 X 2 X 2) possible combinations that may be demanded by a task.

When a new task arrives and the pilot is idle, the pilot will immediately begin work on that task. If the pilot is not idle, but sufficient pilot resources are available, the pilot will also begin work on the second task immediately. If resources are not available, the model will preempt a lower-priority task, if possible. If preemption is not possible, the new task waits until resources become available so that the pilot can begin work on this new task.

Within the simulation, the availability of resources is determined in the following manner. A 2 X 3 matrix maintains a listing of resources currently required by tasks on which the pilot is working. This matrix is shown in Table 1. All values in this matrix are initially set equal to zero. When the first task arrives, the resource requirements of that task are used to change corresponding matrix cell values to equal one, while the others remain equal to zero. If a second task arrives prior to the completion of the first, the resource requirements of the new task are temporarily added to those of the old task. If this results in no more than two resources being used simultaneously, the second task is assigned to the pilot to work. If three resources are required simultaneously, a further computation is required to determine if both tasks can be accomplished simultaneously. This computation is the determination of a Degradation Factor, which is based on the number of overlapping task resource requirements. The scheme used to determine this is shown in Table 2. The values in this table are derived from the conflict matrix developed by North. In the parallel simulation model this Degradation Factor is used both to determine if an additional task can be assigned to the pilot, and to compute the amount of service time degradation resulting from that assignment.

Table 1 : Multiple Resource Model Matrix

MODALITY	CODE	STAGE
Auditory	Spatial	Responding
Visual	Verbal	Encode/Central Processing

Table 2: Degradation Factors

NUMBER OF OVERLAPPING RESOURCES	DEGRADATION FACTOR
0 (With 2 Tasks)	~ Uniform (0.1, 0.3)
1	~ Uniform (0.3, 0.6)
2	~ Uniform (0.6, 0.9)
3	~ Uniform (0.9, 1.1)
4 (Or More)	1.1

Whenever a new task is assigned to the pilot, resulting in the pilot working on more than one task (parallel processing), the service times for all tasks being worked are increased. This reflects the performance degradation resulting from parallel processing, as compared to serial processing, as described by Norman and Bobrow and by Wickens, Mountford, and Schreiner. Service times are increased by computing a Degradation Factor. This factor, normally between 0 and 1, is multiplied by the single-task service time and the product added to that service time. For

example, if a degradation factor of 0.5 were computed, all affected task service times would be multiplied by a factor of 1.5, or effectively increased by 50%. If preemption is not possible and the Degradation Factor exceeds 1.0, then the new task would be assigned to wait until resources are available. A Degradation Factor limit of 1.0 is a subjective, arbitrary limit. It represents doubling a service time and assumes that it is better to defer a task to an idle period than to double the current service times of all tasks being worked at that time. Each time the pilot begins work on a new task, the resources being used to work on all tasks are updated. Similarly, when a task is preempted or completed, resource status is also updated. Figures 3 and 4 display the task arrival and service completion processes respectively for the parallel processing model.

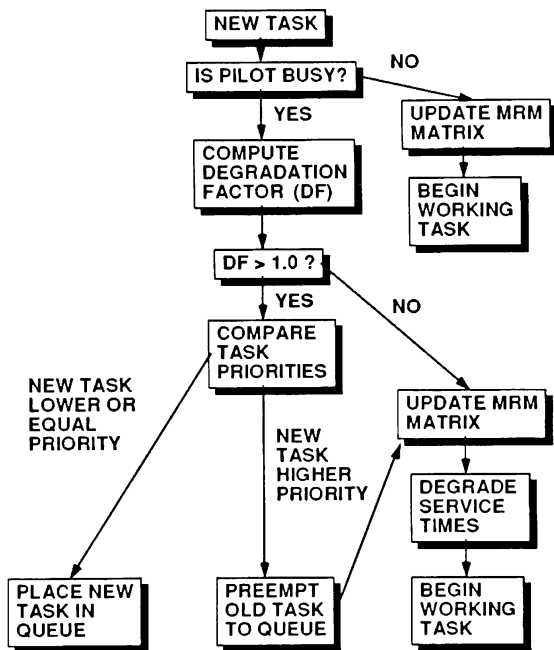


Figure 3: Task Arrival Process, Parallel Processing Model

The method used to degrade service times is based on simply counting the number of required resource "overlaps" and computing a Degradation Factor based on this count. This procedure is a modeling simplification since it ignores some existing evidence presented by Norman & Bobrow and Wickens, Mountford, and Schreiner that the type of resource overlap is important in determining the actual degradation of service times for dual task performance over single task performance. Very little data exists to accurately determine these degradations for specific task combinations if resource type were considered. However, the degradations used were based on North's work and agree with the degradation factors he presented. Therefore, since data for specific resource overlaps was not available, an assumption was made that simply counting the number of required resource overlaps was a reasonable modeling approach.

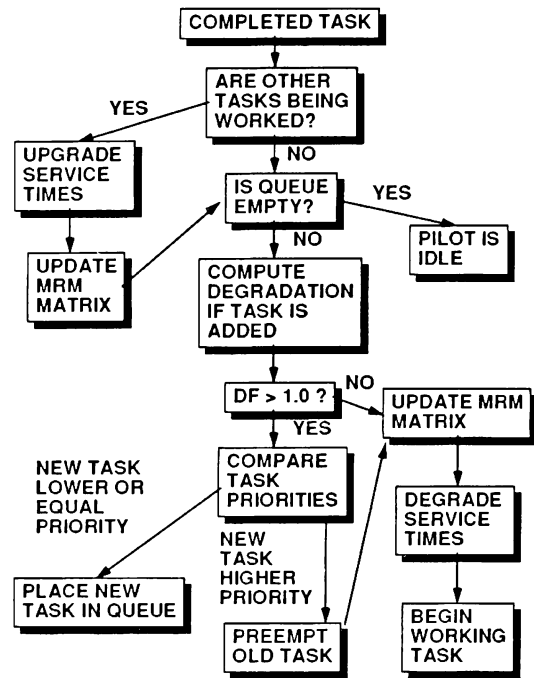


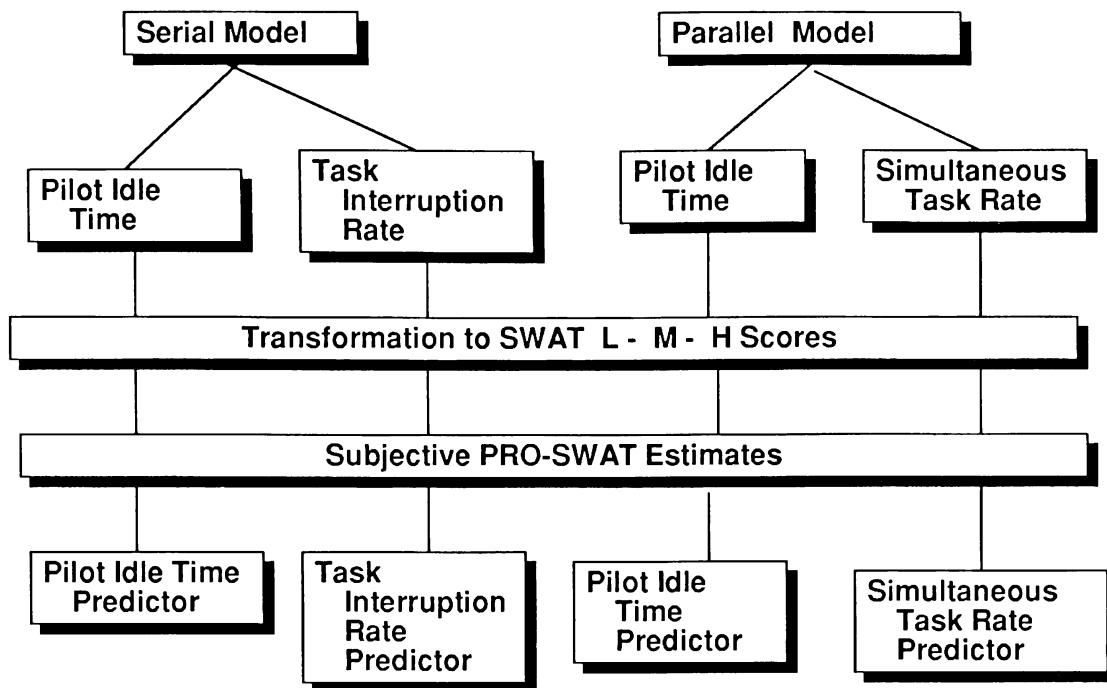
Figure 4: Service Completion Process, Parallel Processing Model

### 3. ANALYSIS AND RESULTS

The previous sections described two modeling approaches used to represent a pilot flying an aircraft. Each of these models takes a fundamentally different approach to how tasks are processed by the pilot. An obvious question arises as to which model more accurately represents reality. This was approached by using the simulation models to predict pilot workload and comparing these predictions to workload measurements taken in a flight simulator. These workload measurements were taken in a U.S. Air Force study, the Radar-Aided Mission/Aircrew Capability Exploration (RAM/ACE) Experiment, by Quinn, Jauer, and Summers.

#### 3.1 Predictors

Seven measures derived from the simulation models were used as candidate predictors of workload. Five of these candidate predictors were based on the serial processing model. These five candidate predictors used pilot idle time and four different definitions of task interruption rate, determined from the serial processing pilot model, as estimators of workload in the Time Load dimension of the SWAT methodology. The remaining two candidate predictors were derived from the parallel processing model of a pilot. These two predictors used pilot idle time and the simultaneous task rate as estimates of workload in the Time Load dimension. Figure 5 displays the structure and relationships of these terms and measures.



**Figure 5: Derivation of Time Load Predictor**

By using pilot idle time derived from both the serial and parallel processing models, conclusions could be drawn about the appropriateness of each model. Pilot idle time or spare time is simply a measure of the proportion of time the pilot is not occupied performing a required task. It does not recognize other factors which may impact in the Time Load dimension. These factors would include the influence of tasks interrupting one-another or the influence of tasks being accomplished simultaneously.

In this serial processing model, only one task can be accomplished at a time, with higher priority tasks interrupting lower priority tasks. Therefore, the relative rate at which this happens, the task interruption rate, measures the effect of this phenomena on the SWAT Time Load dimension. A serial processing model does not allow for simultaneous processing of more than one task, so a simultaneous task rate would be irrelevant for the serial processing model.

For the parallel processing model, tasks are not often interrupted, but additional tasks may be added to a task currently being worked and processed simultaneously. By measuring the rate at which tasks are completed simultaneously, it is possible to capture this aspect of the Time Load dimension. In a sense, this measure goes beyond the concept of idleness or, conversely, "busyness". Simultaneous task rate suggests that the pilot is not only busy, but is occupied with a given number of parallel tasks, on the average. Although some tasks may be interrupted in a parallel processing model, the effect of these interruptions is assumed to be insignificant

compared to the number of tasks completed simultaneously.

### 3.2 Experimental Conditions

The original RAM/ACE flight simulator exercise was conducted under a wide variety of equipment enhancements, mission scenarios, crew size, and both day and night conditions. A subset of these conditions was selected to provide the basis for comparing the seven predictors previously mentioned. Three cockpit equipment configurations were selected. These were the F-15C Advanced Baseline (ABL), the ABL plus Terrain Following/ Terrain Avoidance (TF/TA), and the ABL plus Terrain/Threat Avoidance (TTA). This eliminated four other possible enhancements from consideration, but the three selected appeared to span the range of possible pilot workload. The ABL represents the enhancement with the fewest automatic features and suggests the possibility of high workload. The TF/TA system automates the problem of avoiding contact with the ground, somewhat alleviating the pilot's workload. The TTA system automates both the terrain avoidance problem and the hostile threat avoidance problem. By examining all three of these configurations, situations of high, medium, and low workload would be examined.

Data was available for both day and night missions in the RAM/ACE study. However, only data from day missions was used. RAM/ACE data was available for both one-man and two-man crews, but only the data for one-man crews was used. It was not the intent of this study to model

the complex interaction process that occurs in a two-man cockpit. Several mission scenarios were also defined for each equipment enhancement. The primary difference in these scenarios was the final course flown to the target and the relative location of turn points along the route of flight. The data for different scenarios was combined and treated as being homogeneous. This provided more RAM/ACE data for comparison purposes.

### 3.3 Methods of Comparison

A fundamental issue is how well workload predictions derived from the two simulation models matched workload measurements taken from the RAM/ACE workload measurements. This comparison was accomplished by viewing the predictions and measurements as two multivariate sets in a two-way experiment. This experiment is depicted in Table 3. This table shows that one of two sets of workload predictions are compared to workload measurements for all three cockpit configurations (ABL, TF/TA, and TTA). The data points themselves are triplets, i.e., workload values for the Penetration (P), Target Acquisition (A), and Weapon Delivery (W) segments for each combination of cockpit configuration and method. When viewed this way, the multivariate nature of the data is more obvious. Two multivariate methods were used to compare predictions and measurements: multivariate comparison of means and profile analysis, as described by Johnson and Wichern.

Table 3: Multivariate Comparison Process

METHOD	COCKPIT CONFIGURATION		
	ABL	TF/TA	TTA
Prediction	P/A/W	P/A/W	P/A/W
Measurement	P/A/W	P/A/W	P/A/W

The multivariate comparison of means involved a hypothesis test which compared a 3 X 3 matrix of mean values of SWAT workload measurements to a similar 3 X 3 matrix of workload predictions. One dimension of these matrices represents the flight phase (P, A, or W) while the other dimension represents the cockpit configuration (ABL, TF/TA, or TTA). The multivariate comparison involved examining one predictor at a time and comparing it against the RAM/ACE workload measurement. Since there were a total of eight different predictors for two different predictor data sets, there were a total of sixteen hypothesis tests. Table 4 shows the results of these comparisons for each model, using pilot idle time as a predictor. This table also shows that two sets of SWAT predictions were available and both were used.

The second multivariate technique used to evaluate the predictors was profile analysis. This technique tests the same hypothesis as the multivariate comparison of means.

This test is accomplished in a stagewise fashion. The first stage asks the question, "Are the profiles parallel?", and involves the formulation of an auxiliary hypothesis. The second stage asks the question, "Assuming the profiles are parallel, are they coincident?" Again, an additional auxiliary hypothesis is formed. The third stage of the profile analysis asks the question, "Assuming the profiles are coincident, are the profiles level?", and a third hypothesis is formed. While this third test may be important in some domains, it was not considered important in this case. There is no reason to suspect, for example, that workload in all three flight phases, or all three cockpit configurations, should be the same. In fact, there are many reasons to expect just the opposite to be true. Table 5 shows the results of these comparisons for each model, again using pilot Idle Time as a predictor.

Table 4: Multivariate Comparison of Means, Idle Time Predictor Only

PREDICTOR	DATA SET 1		DATA SET 2	
	F	P > F	F	P > F
Serial Model	0.32	0.81	2.03	0.13
Parallel Model	0.44	0.73	2.34	0.09

## 4. CONCLUSIONS

Many conclusions could be drawn from this research but the fundamental issue addressed is the relative merits of a serial processing model versus a parallel processing model. A serial model is simpler, easier to develop, and conceptually easier to understand. Yet the serial model ignores evidence of actual parallel processing. If one is only interested in the most accurate model, which of these two models is better?

Perhaps the best way to answer this question is to compare the accuracy of the two Pilot Idle Time predictors. Both of these have been identified as providing accurate predictions and they were computed using identical methods in each model. Therefore, it seems logical that differences in accuracy of these predictors should reflect on the models themselves, rather than other factors.

The results of the multivariate comparison of means, as shown in Table 4, show the serial model produced more accurate predictions for both predictor data sets. The parallel model produced slightly worse results for Data Set 1 and worse results for Data Set 2. If a choice is to be made, Table 4 suggests that the serial processing model, using pilot idle time as a surrogate, produces slightly more accurate predictions than the parallel processing model. However, it does not appear that this increase in accuracy is statistically significant. Results from the profile analysis show no trends regarding parallel versus serial models. Examination of the data in Table 5 shows no evidence that either model is more accurate over both data sets and all three flight configurations in terms of being both parallel and coincident. Based on this evidence, no

Table 5: Profile Analysis Results, Pilot Idle Time Predictor  
(F Statistic Shown)

DATA SET	CONFIGURATION	MODEL	PARALLEL?	COINCIDENT?	EQUAL?
1	ABL	Serial	4.63	0.58	10.46
		Parallel	1.77	0.02	8.67
	TF/TA	Serial	1.99	0.79	8.71
		Parallel	0.65	1.29	11.29
	TTA	Serial	0.08	0.24	8.66
		Parallel	0.08	0.24	8.66
2	ABL	Serial	2.11	0.35	8.17
		Parallel	1.23	0.55	7.08
	TF/TA	Serial	5.50	0.13	1.84
		Parallel	5.83	0.76	4.08
	TTA	Serial	1.62	0.04	6.89
		Parallel	1.62	0.04	6.89

strong statement can be made suggesting that either model consistently produces more accurate results. Since the serial processing model was easier to develop, it may be the model of choice.

## REFERENCES

- Battiste, V. and Hart, S. A. (1985). Predicted Versus Experienced Workload and Performance on a Supervisory Control Task. In: *Proceedings of the Third Symposium on Aviation Psychology* (R. Jensen ed.). Columbus, Ohio, 255-262.
- Johnson, R. A. and Wichern, D. W. (1982). *Applied Multivariate Statistical Analysis*. Prentice-Hall, Inc., Englewood Cliffs, N.J.
- Norman, D. A. and Bobrow, D. G. (1975). On Data-limited and Resource-limited Processes. *Cognitive Psychology* 7, 44-64.
- North, R. A. (1985). A Workload Index for Iterative Crewstation Evaluation. In: *IEEE 1985 National Aerospace and Electronics Conference*. Dayton, Ohio, 868-872.
- Potter, S. S. and Acton W. H. (1985). Relative Contributions of SWAT Dimensions to Overall Subjective Workload Ratings. In: *Proceedings of the Third Symposium on Aviation Psychology* (R. Jensen, ed.). Columbus, Ohio, 231-238.
- Quinn, T. J., Jauer, R. A., and Summers, P. I. (1982). *Radar Aided Mission/Aircrew Capability Exploration (RAM/ACE) Task II, Synthesis*. Aerospace Medical Research Laboratory, AFAMRL-TR-82-91, Wright-Patterson AFB, Ohio.
- Reid, G. B., Shingledecker, C. A. Nygren, T. E., Eggemeier, F. T. (1981). Development of Multidimensional Subjective Measures of Workload. In: *IEEE 1981 Proceedings of the International Conference on Cybernetics and Society*. New York, 403-406.
- Schuppe, T. F. (1988). A Methodology for Predicting Pilot Workload. PhD Dissertation, Department of Industrial and Systems Engineering, The Ohio State University, Columbus, Ohio.
- Wickens, C. D. (1984). *Engineering Psychology and Human Performance*. Charles E. Merrill Publishing Company, Columbus, Ohio.
- Wickens, C. D., Mountford, S. J., and Schreiner, W. (1981). Multiple Resources, Task-Hemispheric Integrity, and Individual Differences in Time-Sharing. *Human Factors* 23, No. 2, 211-229.

## AUTHOR'S BIOGRAPHY

THOMAS F. SCHUPPE is an assistant professor of operations research and Head, Department of Operational Sciences, Air Force Institute of Technology, Wright-Patterson AFB, Ohio. He received a B.S. in mechanical engineering from the University of Wisconsin in 1967, an M.S. in systems engineering from the Air Force Institute of Technology in 1976, and a PhD in operations research from The Ohio State University in 1988. He currently teaches several simulation courses as well as courses in systems analysis and operations research. His primary research area is human performance modeling. He is currently a member of ORSA, AFA, and Theta Tau.

Thomas F. Schuppe  
AFIT/ENS  
Wright-Patterson AFB, OH 45433 USA  
(513) 255-3362