Principles of Performance Monitoring with Application to Automatic Landing

J. M. SMITH,* P. B. SCHOONMAKER,† E. E. PYRON,‡ AND R. L. BENBOW§
McDonnell Douglas Astronautics Company-East, St. Louis, Mo.

A new development in the field of control systems is the Performance Monitor, a subsystem which is intended to assess the total (internal) and external state of a controlled dynamic system, and the expected effect of this state upon the safety of the system's operation. Although the concept appears to be of wide applicability, the only application studied by the authors is to aircraft automatic landing systems. In this application, performance monitoring will provide levels of safety and economy of utilization beyond the level possible by the traditional means of redundancy and internal monitoring alone. This paper presents a brief survey of the principles of autoland performance monitoring—from basic mathematics through conceptual design and testing. Hypothetical numerical examples are provided to clarify the concepts presented.

I. Introduction

THE early stages of autoland development saw extensive use of the conventional approach to guaranteeing safety and utility of the control system—reduction of system failures by use of reliable components, redundancy, and "internal" monitoring of subsystems (in-line and comparison monitoring). Although failure prevention and detection are important in themselves, this approach alone has not proved entirely satisfactory as an economical means for reduction of risk; e.g., redundancy has brought its own problems, such as equalization and nuisance disconnects. Even though effective in preventing and detecting failures, the conventional techniques provide no help for the many possible situations (e.g., ILS beam bends and noise, wind shears) which can lead to loss of an aircraft and its passengers without internal failure.

Over the present and future commercial operational spectrum (categories II and III), these difficult situations will in fact occur with non-negligible frequency. Furthermore, in many of these situations the pilot (the only external monitor in the loop) will be lacking in high-fidelity out-the-window information, and simultaneously saturated with low-fidelity information, requiring much scanning, interpreting and cross-checking of instruments. The situation becomes increasingly difficult as touchdown approaches, since the time-to-go approaches, and finally becomes less than, the total of decision and response time.

A Performance Monitor (PM) is a device which makes an assessment of the total (internal and external) state of the aircraft and the landing maneuver, and provides an effective man/machine interface to display this assessment to the pilot. The performance monitor assists the pilot in making go/no-go decisions under pressure and in the face of uncertainty. The goal in a performance monitor design is to substantially reduce landing risk without imposing an unacceptable economic penalty in the number of aborted approaches.

Presented as Paper 71-958 at the AIAA Guidance, Control and Flight Mechanics Conference, Hempstead, New York, August 16–18, 1971; submitted September 13, 1971; revision received January 10, 1972. The authors are pleased to acknowledge their debt to H. L. Harenberg (who first advanced the concept of a Performance Monitor), and to the members of his staff in the Flight Guidance and Controls Department of Douglas Aircraft Company, who developed the performance monitor presently in service on the McDonnell Douglas DC-10.

Index categories: Aircraft Flight Operations; Safety; Navigation, Control, and Guidance Theory.

* Group Engineer, Design, Department E451.
† Group Engineer, Dynamics, Department E933. Member AIAA.
‡ Senior Engineer, Dynamics, Department E933. Member AIAA.
§ Engineer, Dynamics, Department E933.

A. Monitor's Place in the Total System

The information flow in an autoland is depicted schematically in Fig. 1; the switch on the monitor input symbolically allows us to use a single diagram for systems both with and without performance monitors. When the performance monitor is included in the total system, it provides a performance assessment (new input) to the pilot. It is common in the literature to refer to the monitor's performance assessment as "independent," although the notion of independence is usually left undefined.[1–3] The authors have identified three distinct aspects of monitor independence: 1) Hardware independence, in the sense that the monitor's failures occur independently of failures in the control system. 2) Processing independence, in the sense that the monitor "thinks" in a somewhat different way from the pilot, and comes up with a different type of performance assessment. As seen by the pilot, this is a new source of "information"—an independent performance assessment. 3) The highest possible level of monitor independence is informational independence, which can occur only when the monitor has sensors of a different type from the sensors used for the control system (not just replicated sensors of the same type).

B. Wider Applications of Performance Monitoring

This paper is concerned with the application of performance monitoring to the autoland problem. However, it should be evident that the basic concepts are applicable to any dynamic system in which a premium is placed on timely and accurate

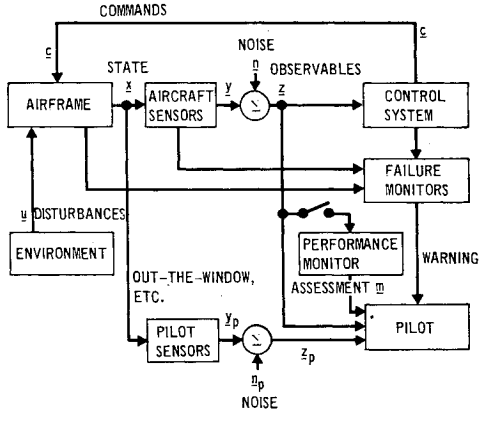


Fig. 1 Schematic of autoland decision process, with and without performance monitoring.

abort/proceed decisions. Such systems have in common the following properties: 1) Dangerous situations can develop "rapidly" (relative to the control response time). 2) Serious safety risks result from proceeding when one should abort. 3) Economic and safety penalties result from aborting when one should proceed.

Systems in which we would expect to find performance monitoring useful, and perhaps absolutely necessary, are: high-speed trains, "flying" monorails and automated high-speed expressways, automatically landed aircraft of all sizes and types (an inexpensive monitor for light planes represents a major design challenge), large ships, which have response time on the order of an hour, and fixed installations such as automated nuclear reactors and oil refineries.

II. Mathematical Developments

Comparing monitor concepts or analyzing the pilot-monitor interaction requires a mathematical model for evaluating landing risk. This section will develop the mathematical model and language in which to express the characteristics of a performance monitor, and the way these characteristics contribute to the overall objective of reducing risk.

A. Abort-Continue Decision Process

The purpose of the monitor is to assist the pilot in making a binary decision: to abort or continue with the landing. Let us define the binary variable p; in symbols we will say $p = C$ or $p = A$, to indicate a pilot decision to continue or abort, respectively.

To analyze the over-all decision process, it is accordingly necessary to discretize the entire process. Let us assume for the moment, then, that the PM also has only a single binary output, defined by the variable m. We say either $m = D$ if the monitor recommends continuation of the approach, by remaining dark, or $m = L$ if the monitor recommends aborting, by lighting up. (We shall later discuss other possible monitor outputs.)

Finally, to discretize the real world, it is necessary to take the continuum of possible state variables at a particular instant, and partition it by absolutely unambiguous criteria into a region G, the set of all "good" approaches and landings, and the complementary region B, the set of all "bad" landings. Figure 2 indicates the external state variables entering into classification at the instant of touchdown. Internal variables, such as the condition of the landing gear and brakes, are also involved in classifying a landing. Corresponding to this classification, we define the binary variable a, and say $a = G$ or $a = B$, according as the approach is good or bad. For convenience, the definitions of these variables are collected in Table 1.

The final notational device needed to proceed with the analysis is the "decision tree". A tree represents a stagewise decision process as a sheaf of diverging paths representing all alternative possibilities at the first stage. Each such path

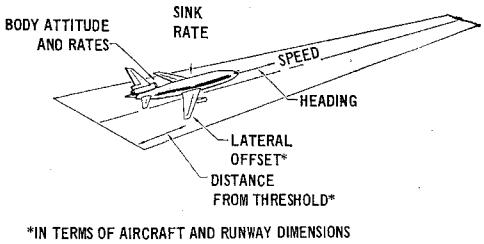


Fig. 2 External state variables involved in good/bad classification at the instant of touchdown.

Table 1 Binary variables used in analyzing the decision process

Name	Definition	Possible values and their meaning
p	Pilot decision	$p = C$: pilot decides to continue approach $p = A$: pilot decides to abort approach
m	PM assessment of approach	$m = D$: monitor dark; recommends continuing $m = L$: monitor lights; recommends aborting
a	True status of approach	$a = G$: approach will culminate in good landing $a = B$: approach will culminate in bad landing

terminates in a node from which emanate paths denoting the possibilities at the second stage, given the decision made at the first stage, and so forth until all possibilities at every decision stage have been exhausted. The final set of nodes encompasses all possible eventual outcomes of the process.

At each node, we mark each outward path with the (known or assumed) conditional probability that, given that the process has arrived at that node, it will then take that path. These conditional probabilities are, of course, "locally exhaustive"; i.e., they encompass all locally possible occurrences and thus sum to 1.0 at every node. The over-all (nonconditional) probability that the process will eventually arrive at a particular final node is found by traversing the complete path from the initial node to that final node, applying the multiplication rule at every node encountered along the way. The final nodes are "globally exhaustive"; thus, the sum over all final nodes is also 1.0.

B. Operational Risk and Penalty

Consider a pilot making an automatic approach and landing with no performance monitor in the loop (i.e., the switch shown in Fig. 1 is open). There is a finite probability that the approach is actually bad. The pilot, in evaluating the approach and making his abort/continue decision, has finite probabilities of assessing a good approach to be bad, and vice versa. These two types of errors have common names in the literature[4]; using conditional probability notation, we have the definitions¶

$$\alpha_p = P(C|B) \qquad \beta_p = P(A|G)$$

Figure 3 is a tree showing the structure of the unmonitored autoland decision process.** Hypothetical numerical values

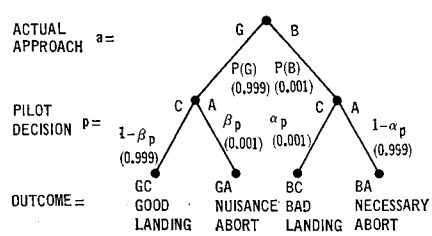


Fig. 3 Tree showing unmonitored autoland decision process.

¶ The subscript "p" denoting a pilot decision.

** These are "one-shot" probabilities, which do not account for repeated tries by those aircraft which abort their initial approaches.

Table 2 Over-all probabilities of all possible final outcomes of the unmonitored autoland decision process[a]

	Good approaches	Bad approaches
Completed	$(1-\beta_p)P(G)$ (0.998 001)	$\alpha_p P(B)$ (0.000 001)
Aborted	$\beta_p P(G)$ (0.000 999)	$(1-\alpha_p)P(B)$ (0.000 999)

[a] Hypothetical numerical values, $P(B) = \alpha_p = \beta_p = 0.001$.

have been given for each of the conditional probabilities involved; these values are purely illustrative, and not assumed to characterize actual autoland systems or pilots. The over-all probabilities of each final result are given in Table 2. The diagonal entries of this table represent the results of correct decisions, and the off-diagonal entries are the results of incorrect decisions:
$\alpha P(B)$ is the over-all total system risk of executing a bad landing, $\beta P(G)$ is the penalty (nuisance-abort probability) associated with the total system.

The risk probability is dependent upon both the autoland performance and the pilot's decision-error performance, while the penalty is almost entirely due to the pilot (except insofar as the pilot's experience with the autoland system influences his decision). Assume that we want to reduce the risk/penalty probabilities without changing the aircraft systems; we then demand lower decision-making errors from the pilot. The given example happened to have $\alpha = \beta$. Generally, a process with high α and low β is called optimistic, one with high β and low α is called pessimistic or conservative.

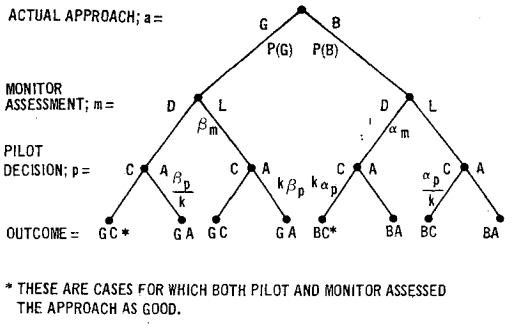
We might expect that, without improving the pilot's working environment, we cannot reduce both his error probabilities simultaneously. We may find pilots who are more or less cautious (alternatively, have higher or lower tolerance for apparent anomalous behavior), but changing the degree of caution will simply trade one kind of error for the other.

C. Influence of the Monitor

Now let us close the switch shown in Fig. 1, and allow the PM to make assessments and display them to the pilot. The tree shown in Fig. 4 shows the over-all decision structure of this process. The means which we use here to model the influence which the PM's assessment exerts upon the pilot's decision-making is a set of (one or more) "influence coefficients" which express the pilot's tendency to alter his inherent error probabilities in the direction of agreement with the monitor's assessment. In the example shown, we use only a single coefficient, k, whose value is 10.0; i.e., the pilot's error probability is decreased by a factor of 10 whenever the monitor is correct, and increased by a factor of 10 when the monitor is incorrect; e.g.,

$$P(A/GD) = 0.1\beta_p = 0.0001 \qquad P(A/GL) = 10\beta_p = 0.01$$

The monitor's assessment errors are denoted α_m, β_m.



* THESE ARE CASES FOR WHICH BOTH PILOT AND MONITOR ASSESSED THE APPROACH AS GOOD.

Fig. 4 Tree showing effect of monitor upon autoland decision process.

Final result probabilities are shown in Table 3. For the example shown, substantial reductions in both risk and penalty have been achieved, as compared to the unmonitored autoland of the preceding example. Although this precise result is a function of the numbers used, the general result (that the monitor acts to reduce both risk and penalty) is very insensitive to the numerical values of α_m, β_m and k; e.g., with k fixed at 10.0, it holds true up to $\alpha_m = \beta_m = 0.091$; alternately, holding $\alpha_m = \beta_m = 0.001$, it remains true for $0 < k < 999$. In short, it is easy to demonstrate, from a remote vantage point such as provided by the decision-structure analysis just given, that the pilot should side with the monitor almost all the time.

Consider, however, the view from the cockpit; suppose the monitor lights, but the pilot proceeds to landing anyway; what is likely to happen? For reasonable levels of control system and monitor performance, the monitor will light up in only a small percentage of all approaches. Some of these lightups will be actual bad approaches, while others will be false alarms. Since both of these probabilities are small and variable numbers, it is evident that their ratio can vary over a wide range. We can provide an equation for this ratio by using Bayes' theorem[5]:

$$P(G/L) = \frac{P(L|G)P(G)}{P(L)} = \frac{P(L|G)P(G)}{P(L|G)P(G) + P(L|B)P(B)}$$

For the example given, this probability is 0.5; i.e., out of all cases in which the monitor lights up (two approaches per thousand), exactly half are false alarms! This fact could in some individual cases have the effect of "spoofing" pilots; i.e., giving a mistaken impression of the magnitude of the monitor's β error, with potentially serious consequences at some later time.

This possibility could be prevented by establishing an operational ground rule calling for a mandatory abort whenever the monitor lights. No tree is provided to show the decision structure for this case; however, referring back to Fig. 4, you will see that only the two paths marked with asterisks will lead to completed landings under this rule, and all others will be aborted. The final outcome probabilities, given in Table 4,

Table 3 Monitor's influence upon probabilities of final outcomes of autoland decision process[a]

	Good approaches	Bad approaches
Completed	$\left[(1-\beta_m)\left(1-\frac{\beta_p}{k}\right)+\beta_m(1-k\beta_p)\right]P(G)$	$\left[\alpha_m k\alpha_p+(1-\alpha_m)\frac{\alpha_p}{k}\right]P(B)$
		$(0.1\times10^{-6}$ vs $0.1\times10^{-5})$
Aborted	$\left[(1-\beta_m)\frac{\beta_p}{k}+\beta_m k\beta_p\right]P(G)$	$\left[\alpha_m(1-k\alpha_p)+(1-\alpha_m)\left(1-\frac{\alpha_p}{k}\right)\right]P(B)$
	$(0.1\times10^{-3}$ vs $0.999\times10^{-3})$	

[a] With hypothetical numerical values, $P(B) = \alpha_p = \beta_p = \alpha_m = \beta_m = 0.001$, $k = 10$.

Table 4 Probabilities of final outcomes under mandatory—abort rule[a]

	Good approaches	Bad approaches
Completed	$(1-\beta_m)\left(1-\frac{\beta_p}{k}\right)P(G)$	$k\alpha_m\alpha_p P(B)$ (0.1×10^{-7})
Aborted	$\left[(1-\beta_m)\frac{\beta_p}{k}+\beta_m\right]P(G)$ (0.0011)	$[\alpha_m(1-k\alpha_p)+(1-\alpha_m)]P(B)$

[a] With hypothetical numerical values as before.

are what we should expect from previous discussions—risk is decreased only by increasing penalty. One can also imagine that a mandatory-abort rule might be unpopular with pilots, especially in CAVU.‡‡

The proper design approach to reducing the likelihood of pilot override of the PM is a) engineering the monitor for low decision errors, to justify pilot confidence, and b) human-engineering the monitor to build man/machine "rapport," and establish pilot confidence.

D. Additional Monitor Outputs

Part of this human-engineering task is to provide the pilot with PM outputs other than the simple dark or light alternatives considered thus far. The purpose of such outputs is to supply the pilot with more detailed data regarding the monitor's assessment.

Discrete outputs from the monitor may offer the pilot three or more alternatives; e.g., continue landing on autopilot, takeover and land manually, abort on autopilot, takeover and abort manually. (The extra alternatives correspond to an attempt by the PM to characterize the difficulty of the present situation as being internal or external in origin.)

Continuous outputs, however, offer the best potential for establishing and maintaining man/machine rapport. Continuous outputs may be classed as either "envelope" or "point" signals. An example of each type is given in Fig. 5: the cross provides a point signal, the ellipse an envelope signal.

Envelope signals pictorially indicate (by means of a display which can grow and shrink, change color or whatever) the monitor's estimate of the uncertainty or dispersion in the present or predicted state. This gives the pilot insight into the status of and reason for the monitor's current assessment, which he can continuously weigh against his own perceptions. The abort decision process then becomes one of a gradual deterioration in confidence, rather than a sudden crisis of indecision.

Point signals allow the monitor to indicate not just how well it estimates the approach to be going, but the magnitude and specific direction by which it estimates the present or predicted state differs from the nominal. As long as no takeover occurs, the effect of such signals is not much different from envelope signals. In the event of a takeover, however—especially in the first critical seconds—the pilot may be strongly included to steer in the direction favorably indicated by the monitor. This blurs the distinction between monitor and control system; it is a bit difficult philosophically to reconcile this post-takeover secondary control function with the pretakeover monitor/control system independence. In some circumstances, it may be appropriate to blank all point signals at the onset of a takeover or abort.

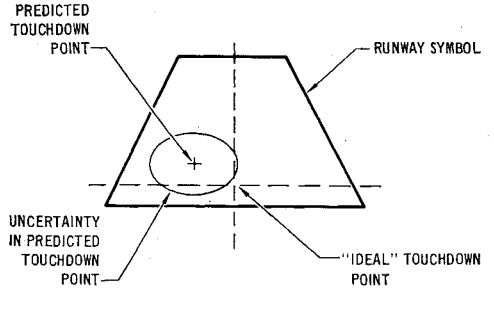


Fig. 5 Example of continuous monitor displays.

‡‡ Ceiling and visibility unlimited.

III. Key Issues in Performance Monitor Design

We have seen that the intended function of a performance monitor is to provide the flight crew with an autoland performance assessment (which is independent of the crew's assessment), as the aircraft and autopilot work together to execute a safe landing. It is the designer's task to realize that goal within economic constraints. This task includes selection of data-handling, computational and display hardware; design of computational algorithms; and finally, selection of numerical values for the parameters of the algorithms, to maximize their effectiveness in the working environment of the particular airplane/autoland system.

Typically, a monitor will have two distinct types of parameters: a) 'Fidelity' parameters which tailor the PM response to the particular system in question. Theoretically, variation of a parameter of this type can reduce both α and β errors simultaneously. b) 'Caution' parameters which establish the departure from postulated ideals which the monitor will tolerate without generating an alarm. Variation of any parameter of this type will drive α and β errors in opposite directions.

Unlike control systems, whose maximum effectiveness is limited by the basic physics of the airplane, the monitor caution parameters can be set to give a bad-landing risk as low as you please, but at the expense of high nuisance-abort penalties. To have an economically viable system (e.g., one which increases traffic flow over current levels under all certifiable weather minima), one must decrease the nuisance-abort penalty to a level corresponding to an acceptable level of landing risk. We see then that ultimately, the PM is an implementation of the design engineer's approach to performance assessment and will embody the designer's judgement as to what constitutes an "acceptable" level of risk (e.g., an order-of-magnitude reduction below current landing-accident rates). In a sense, then, the PM puts a controls engineer in the cockpit.

A. Sources of Data

The performance monitor will normally interface with failure-monitoring devices (in-line and comparison monitors), and with aircraft onboard sensors which convey raw data relating to the aircraft external state, gyros, accelerometers, ILS receivers, etc. To accomplish its objectives, the PM should have raw-data sources *independent* of the control sys-

tem's data sources. Then, if the control system is carrying the aircraft into an intolerable situation because of bad input data received from its sensors, the monitor will have a high probability of detecting this condition, on the basis of data from its independent sensors.

The lowest level of independence providing this probability is hardware independence, in which the monitor's sensors are distinct from, but copies of, the sensors feeding the control system. (This would be easy to accomplish in a "hot spare" redundancy scheme.) In the event that the source of the bad control system data is degradation of one of its sensors, the monitor's sensors will still provide good data.

This level of independence does not cope with bad data whose source is external to the aircraft sensors; e.g., anomalies in the ILS beam as radiated, which will affect all replicated sensors alike. To detect such events, the monitor would need a sensor of a completely different type; e.g., Doppler radar. Thus the highest level of independence is afforded when the PM can base its assessment upon information of a type which is not available to the control system.

This leads to an interesting design dilemma: if some new sensor is effective in monitoring the existing control system, it would probably be even better (from a total system standpoint) to modify the control system to use data from this new sensor! Thus we must conclude that in a well-designed total system, the monitor will have its own replica sensors, but will not have sensors basically different from those used for control.

We might reiterate that a monitor can increase the amount of information made available to the pilot, even though in theory the monitor generates no information beyond what is inherent in the sensor inputs. This is because the monitor can "condense" a large mass of data to emphasize a few essential features, and present these to the pilot in a format which he grasps instantly.

B. Scope of the Performance Assessment Function

The greatest benefits of performance monitoring accrue in monitoring the performance of the controlled system (i.e., goal-directed motion of the entire aircraft) as well as the control system. We thus directly address the essential questions: Is the aircraft under control? Is the automatic landing process converging? Will the pending landing be safe? The broadest possible scope of the monitoring function would involve an attempt to estimate the status of the total system, the controlled system elements and the uncontrolled environment. Difficulties are encountered in trying to extract, from aircraft onboard sensor data, a description of the disturbance environment (e.g., winds, beam anomalies, etc.) to which the aircraft is subjected. However, there are considerable benefits possible if the attempt is successful. This type of information is particularly useful in reducing bad-weather abort penalties: with a good indication of the cause of a prior abort, the pilot can, when appropriate, attempt another approach at his original destination, rather than diverting to a field having better weather.

C. Assessment Criteria

The three basic questions of controllability, convergence, and landing safety are not attacked by the PM as abstracts. Rather, performance criterion-functions or indices of control (performance laws analogous to control laws) are developed as part of the design process. During each landing approach the PM uses onboard observables to compute the performance criterion-function values, and operates upon these to formulate a performance assessment: a) The aircraft is assessed to be under control provided that the indices of controllability are within acceptable bounds (i.e., the performance laws are satisfied). b) The autoland process is assessed to be converging, provided that the values computed for certain settling-time parameters are found to be sufficiently smaller than the time-to-go to touchdown. c) The pending landing is assessed to be safe provided that the predicted touchdown point falls within a specified region on the runway, and the attitude, sink rate, and lateral drift rate are within acceptable bounds. The boundaries of acceptability are established partly by aircraft geometry and structural limitations, and partly by analysis of rollout dynamics.§§

Short-term averaging of the input observables and/or the computed output values is normally used for performance assessment. This averaging is intended to smooth the assessment process, thus reducing display "jitter," without adding significant delays. Typically the averaging period will be reduced as the time-to-go decreases.

D. Implementation

Two scalar measures of performance have been used to address the question of controllability and convergence: 1) settling time, and 2) an index of the degree to which guidance and control laws are satisfied. In this paper, a performance index (PI) or controllability index is a positive definite scalar function of several system states or observables, whose value is to be a measure of the degree to which the aircraft is under control. For example: consider an autopilot whose lateral trajectory control loop is based upon a guidance law of the form

$$\phi_c = a_1 y + a_2 \dot{y} + a_3 \int_{t_0}^{t} y(t)\, dt$$

where ϕ_c is the roll command, y is the lateral displacement from the ILS localizer centerline, the a_i are the control system gains, and t_o is the time the control mode was selected. The degree to which this guidance law is satisfied is measured by a performance index based upon the difference between the monitor estimate of the commanded state $\hat{\phi}_c$ and the actual measured state ϕ_m:

$$I = (\phi_m - a_1 \hat{y} - a_2 \hat{\dot{y}} - a_3 \int_{t_0}^{t} \hat{y}(t)\, dt)^2 \P\P$$

The smoothed value of such a PI may drive a display directly, or may be magnitude gated and used to set performance assessment logical discretes, answering the question "Is the aircraft under control?" Performance indices need not mimic control laws, although there are certain advantages to doing so, as discussed later under Modeling.

The settling time of a control system is usually defined as the time required for the control system to bring the aircraft within acceptable position and velocity bounds relative to the runway. In principle, settling time can be computed from the controlled system Hamiltonian and the rate of change of the Hamiltonian, by the relationship

$$T = \max(H/\dot{H})$$

with an appropriate sign convention.[6] In practice, this method of settling time computation (and some of its variations) has met with little success. A workable approach has been developed, and will be discussed under Modeling. Settling time, like the performance indices, is a scalar. It provides information directly related to the question "Is the landing process converging?"

E. Types of Prediction

Until the aircraft comes to a stop, it is still exposed to risk. Thus, every performance monitor must include a capacity to predict its "end" condition. Prediction may be implemented

§§ In any case, these bounds must fall within those established by contractual or Federal specifications.

¶¶ Note that ϕ_c is the control system command while $\hat{\phi}_c$ is the command the monitor would expect from the control system. The monitor estimate is based on its set of sensors and duplicate control laws.

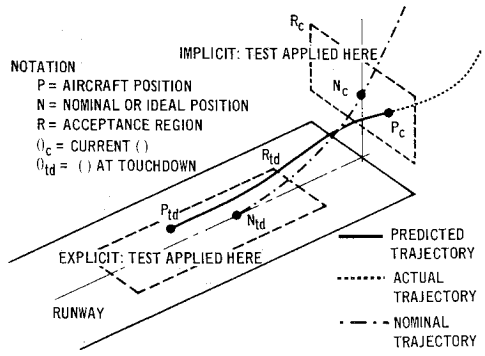


Fig. 6 Implicit and explicit prediction to touchdown.

either explicitly or implicitly; in fact, both approaches may be used in a single PM design, to answer the question, "Will the pending landing be safe?"

In explicit monitoring, performance is assessed by predicting the aircraft state at some critical future event, such as touchdown, and comparing it with the acceptance region of touchdown states (R_{td}). In this way, monitoring is performed explicitly in terms of the final values of the coordinates which the system is trying to control. Implicit monitoring is accomplished by testing the aircraft state at the current time against an acceptance region relevant to the current time (R_c). (Normally the bounds on R_c will converge as the current time advances.) In this method, a safe future state is *implied* by the acceptability of the current state. These two approaches, shown schematically in Fig. 6, are analogous to explicit and implicit guidance mechanizations.[7]

Explicit prediction requires a model of the controlled system dynamics, and implicit prediction does not. Although the model need not be highly accurate, modeling does tend to make explicit prediction more complex to implement than implicit prediction.

Full-explicit monitoring is difficult to implement; e.g., when the aircraft is still several thousand feet from touchdown, it is difficult to predict what its state will be at the end of braking. Therefore, explicit monitoring is at most "phasewise explicit," i.e., explicit for the duration of each approach phase (such as ILS track, flare, and rollout) with implicit prediction used at the phase boundaries. In principle, this is what the FAA's Category 2 rules do, for example, in specifying an acceptance region at the 100-ft decision height in addition to one at touchdown.[8]

F. Role of Modeling

A key issue in monitor implementation is the use of modeling. Unlike control systems, which must be tailored to the controlled-element dynamics, it is not necessary (in principle) to use modeling or other a priori knowledge of the controlled system dynamics to monitor the system's accomplishment of its objectives. For example, monitoring without models can be accomplished by: a) Assessing controllability on the basis of testing high-frequency (inner loop) aircraft motion variables or functions of these variables against constant or time-varying bounds. b) Estimating settling time on the basis of observed displacements and rates. c) Assessing landing safety by testing low-frequency (outer loop) attitude and trajectory displacements against constant or time-varying bounds. Accurate and timely estimation of settling time without models is rather difficult. This is because without a model the aircraft motion is known only to a few derivatives removed from the observables, and the settling time computation is quite sensitive to errors in these derivatives. This demands considerable smoothing of the measured data, and hence a time lag in the assessment. As the landing nears completion, this lag becomes of the order of the time remaining to touchdown.

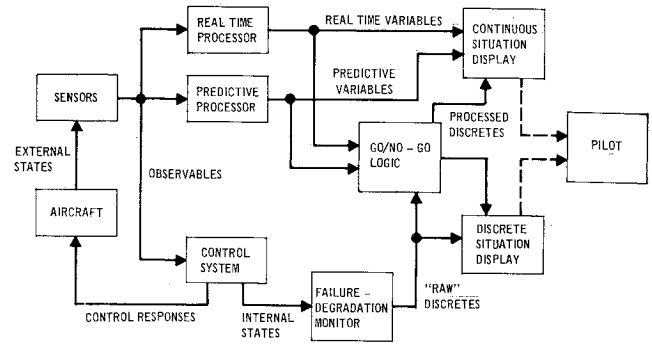


Fig. 7 Over-all monitoring functional flow.

The advantage of monitoring without models is the simplicity of implementation, design, development, and test, as compared to monitors with models. This tends to hold down the development costs of the monitor. The chief disadvantage to monitoring without models is a reduced ability to incorporate a priori information about the system and thus favorably shape the α-vs-β response. In the terminology offered earlier, a monitor with no models has only caution parameters and no fidelity parameters. Thus any reduction in risk is always accompanied by an increase in penalty, and hence higher operating costs.

By comparison, monitoring with models can be accomplished by: a) Assessing controllability on the basis of the degree to which the guidance and control laws are satisfied. b) Estimating settling time from the assumed controlled-system dynamics—either by fast-time simulation, or by auxiliary equations which generate settling time directly. c) Assessing landing safety on the basis of the predicted dispersions at the time of touchdown (or the next upcoming critical event), using means similar to the estimation of settling time.

The principal difficulty with the use of predictive models is initializing the models. Estimating the current state by filtering measured data can lead to a high data-processing load. The advantages are that: a) Modeling provides a straightforward, intuitively defensible means of assessing controllability, convergence, and landing safety. b) There are no irreducible delays in the assessment process, so the assessment time can be made negligible relative to the system response time. c) Nonlinear elements in the control process are easily handled in fast-time simulation models. In short, the more known about the control and controlled system, and the more this knowledge is embodied in the monitor by use of models, the better the monitor performance will be. Partly for this reason monitoring is considered to be part of control system technology.

In summary, Fig. 7 shows the functional blocks which one would expect to find in an automatic landing system performance monitor design. Examples of the detailed transfer functions for each of these blocks are presented and discussed in Ref. 1.

IV. Design Verification

As the monitor design evolves, it becomes necessary to evaluate its performance. (By evaluate we mean to establish with high confidence.) Testing of the monitor in combination with the controlled system will supply the data necessary to compute $P(B)$, α and β, and hence the total-system risk and penalty. The test data will also help in redesign if performance is not totally satisfactory.

A. Verification Problem

The problem arises in attempting to estimate, with high confidence, the probability of occurrence of events which occur with very low probability. Improvements in sampling meth-

odology are necessary to make it feasible for any manufacturer to gather "enough" data to validate his system performance at the desired levels of risk. For example, suppose we wish to verify that the bad-landing probability $P(B)$ is less than one per thousand, by direct sampling. Assume we have a Monte Carlo simulation of the total system (environment, airframe, control system, and monitor). For each Monte Carlo run, a random number generator samples all aircraft disturbances, and the complete approach is simulated. We record the occurrence of good and bad landings. To establish a low bad-landing probability with high confidence, we must have long strings of good landings. For example, 1000 landings can provide at best 63% confidence that $P(B) \leq 0.001$ (i.e., if all 1000 were good; if even one was bad, our confidence that $P(B) \leq 0.001$ would drop to 26%). Similarly, it would take at least 2300 runs to verify $P(B) \leq 0.001$ with 90% confidence, 3000 for 95% confidence, and 4500 for 99% confidence. Current autoland development goals are to limit the undetected bad-landing frequency to something of the order of 10^{-6} to 10^{-8}. Verifying such performance with direct sampling techniques requires an impractical number of runs.

The solution to the verification problem results from the use of "extrapolative techniques", whereby a data base of economically feasible size is transformed into results meaningful for the entire sample space.

B. Stratified Sampling

Stratified sampling techniques have long been used in Monte Carlo analysis[9] and opinion polling.[10] The method, simply stated, is to a) find the conditional probability that a given disturbance will cause a bad landing, regardless of the probability that the disturbance will actually occur, then b) use the known probability density of the disturbance as a weighing function to find the overall probability of a bad landing.

For a single disturbance, u, the basic data to be recorded at various values of u are $P(B/u)$, $P(L/u)$, and $P(BL/u)$, which are respectively the conditional probabilities that the landing is bad, the monitor lights up, and both occur simultaneously, given a particular value of u. Manipulation of the basic data gives the decision-error functions,

$$a(u) = P(BD|u) = P(B|u) - P(BL|u)$$
$$b(u) = P(GL|u) = P(L|u) - P(BL|u)$$

The probability density $p(u)$, is used as a weighing function in the summations

$$P(B) = \sum_u P(B|u)p(u), \qquad P(G) = 1 - P(B)$$

$$\alpha = (1/P(B)) \sum_u a(u)p(u)$$

$$\beta = (1/P(G)) \sum_u b(u)p(u)$$

Assume that the objective is to verify with 95% confidence that $P(B) < 0.001$. One way this can be achieved by direct sampling is to have 3000 good landings without a single unsatisfactory one. This is, of course, the minimum number of runs which will accomplish the objective. The "typical" (median) number of runs required depends strongly upon the "spread" between the true, unknown $P(B)$ and the test value $\hat{P}(B)$. For example, if the true $P(B) = 0.0005$, a typical direct sampling test program will require 9065 runs, with 4 unsatisfactory landings, to establish $P(B) < 0.001$ with 95% confidence. Stratified sampling can be used to reduce the number of runs by an order of magnitude.

Table 5 shows a possible decomposition of the over-all true and test probabilities into three regions, or strata, the regions being defined by the magnitude of the disturbance u. If we now establish 95% confidence for each of the component $P(B/u)$ values, we necessarily obtain 95% confidence for the over-all $P(B)$. In a practical test program, a mix of techniques is used to establish confidence in each region. In region 1, where bad landings are very infrequent, we might estimate $P(B/1)$ by estimating the standard deviation of landing dispersions (from simulation or flight test data), assuming gaussian statistics, and comparing the landing dispersions to the size of the landing acceptance region on the runway. In regions 2 and 3, the bad-landing frequency is high enough to estimate $P(B/u)$ by counting simulated bad landings; it is unnecessary to assume gaussianness of the dispersions. It is found that in a typical 95% confidence test program using stratification, 32 samples are required for region 1, 445 samples for region 2 and 278 samples for region 3, for a total of 755 samples using stratification, as compared with the direct-sampling total of 9065 samples.

In the real world, of course, there are many different disturbance sources. Many of these can be considered independent, but others are highly correlated; e.g., strong winds and water on the runway. Vectorial summations of the form

$$P(B) = \sum_{\mathbf{u}} P(B|\mathbf{u})p(\mathbf{u})$$

and similar forms for α and β, are not difficult to evaluate from results relating to one error disturbance at a time, if the disturbances are independent. Methods for treatment of correlated disturbances by use of covariance techniques[11,12] are still under development.

C. Parameter Extrapolation

The second technique to be discussed is much the opposite of the first; rather than obtaining data from the low-probability critical region, we establish a sequence of artificial critical regions, in each of which direct sampling can economically be applied. This is done by setting all monitor caution parameters to such a low tolerance level that the monitor lights with high probability, say 0.9, on normal approaches. In other words, we force a high β error. When a reliable value has been established for the β error associated with the initial caution parameters, the tolerances are opened up to give a lower β error, say 0.5, and sufficient samples are taken to establish this new level with confidence. Naturally, the second set will take more samples than the first to provide the same confidence. This process is repeated to the extent economically feasible; the available data is then extrapolated to establish the β error expected at the design tolerances.

Table 5 Numerical data illustrating stratified sampling

Region	Known $p(u)$	Unknown true probabilities		Test probabilities	
		$P(B\|u)$	$P(B\|u)p(u)$	$\hat{P}(B\|u)$	$\hat{P}(B\|u)p(u)$
1	0.97	10^{-6}	0.97×10^{-6}	10^{-4}	0.000097
2	0.02	0.01	0.0002	0.02	0.0004
3	0.01	0.03	0.0003	0.05	0.0005
Sum	1.00	—	0.0005	—	0.001

This method is suitable for flight test, as well as simulation, since it does not require exposing the aircraft to bad-landing risk. It is intended primarily for the evaluation of β error. However, it should be clear that the two techniques can be used in combination, and in combination with other techniques, such as sequential sampling[13], to further improve the economics of the design verification program.

V. Summary

The performance monitor is a new subsystem in automatic control. It is specifically related to the safety of automatic operations when marginal control situations can frequently be encountered. The approach the authors have developed for modeling and analyzing performance monitor risk and penalty and their interrelationship should be useful in monitor trade studies, monitor concept selection, and identifying monitor system operating ground rules. The discussion on independence, scope, performance assessment criteria, types of monitors, and modeling should clarify issues that have been quite confusing. The analogy between monitor performance indices and control laws should help the controls engineer identify with the key monitor design task of synthesizing "performance laws" which, when satisfied, indicate satisfactory performance. Finally, the discussion of test and verification should provide monitor designers with our approach to answering the question "How does one economically verify, with high confidence, an ultra-reliable monitored control system?"

References

[1] Harenberg, H. L., Jr. and Shannon, J. H., "Development of the DC10 Automatic Landing Monitor," Paper 690672, Oct. 1969, Society of Automotive Engineers, Los Angeles, Calif.

[2] Parks, D. L., and Tubb, D. G., "Simulator Development of a Perspective Display as an Independent Landing Monitor," AIAA Paper 70-924, Los Angeles, Calif., July 1970.

[3] Bechtel, B., "Radar Independent Landing Monitors," AIAA Paper 70-1336, Houston, Texas, Oct. 1970.

[4] Hogg, R. V. and Craig, A. T., *Introduction to Mathematical Statistics*, Macmillan, New York, 1965.

[5] Parzen, E., *Modern Probability Theory and Its Applications*, Wiley, New York, 1960.

[6] Schultz, D. G. and Melsa, J. L., *State Vector Control Theory*, McGraw-Hill, New York, 1969.

[7] Pitman, G. R., ed., *Inertial Guidance*, Wiley, New York, 1962.

[8] "Criteria for Approving Category 1 and Category 2 Landing Minima for FAR-121 Operators," FAA Advisory Circular 120–29, Sept. 1970, Washington D.C.

[9] Kahn, H., "Use of Different Monte Carlo Sampling Methods," *Symposium on Monte Carlo Methods*, Wiley, New York, 1956.

[10] Deming, W. E., *Some Theory of Sampling*, Dover, New York, 1966.

[11] Noton, A. R. M., "The Statistical Analysis of Space Guidance Systems," JPL TM 35–15, June 1960, Jet Propulsion Laboratory (California Institute of Technology), Pasadena, Calif.

[12] Merel, M. H. and Mullin, F. J., "Analytic Monte Carlo Error Analysis," *Journal of Spacecraft and Rockets*, Vol. 5, No. 11, Nov. 1968, pp. 1304–1308.

[13] Boot, J. C. G., *Mathematical Reasoning in Economics and Management Science*, Prentice-Hall, Englewood Cliffs, N.J., 1967.