PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

Phil. Trans. R. Soc. A (2007) **365**, 623–632
doi:10.1098/rsta.2006.1927
Published online 12 December 2006

Damage prognosis: the future of structural health monitoring

By Charles R. Farrar[1,*] and Nick A. J. Lieven[2]

[1]*Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[2]*Department of Aerospace Engineering, University of Bristol, Bristol BS8 1TR, UK*

This paper concludes the theme issue on structural health monitoring (SHM) by discussing the concept of damage prognosis (DP). DP attempts to forecast system performance by assessing the current damage state of the system (i.e. SHM), estimating the future loading environments for that system, and predicting through simulation and past experience the remaining useful life of the system. The successful development of a DP capability will require the further development and integration of many technology areas including both measurement/processing/telemetry hardware and a variety of deterministic and probabilistic predictive modelling capabilities, as well as the ability to quantify the uncertainty in these predictions. The multidisciplinary and challenging nature of the DP problem, its current embryonic state of development, and its tremendous potential for life-safety and economic benefits qualify DP as a 'grand challenge' problem for engineers in the twenty-first century.

Keywords: damage detection; prognosis; life prediction

1. Introduction

As structural health monitoring (SHM) technology evolves and matures, it will be integrated into a more comprehensive process referred to a *Damage Prognosis* (DP). It is defined as the estimate of an engineered system's remaining useful life (Farrar *et al.* 2003). This estimate is based on the output of models that develop behavioural predictions: by coupling information from usage monitoring; SHM; past, current and anticipated future environmental and operational conditions; the original design assumptions regarding loading and operational environments; and previous component and system level testing and maintenance. Also, 'softer' information such as user 'feel' for how the system is responding will be used to the greatest extent possible when developing DP solutions. In other words, DP attempts to forecast system performance by measuring the current state of the system (i.e. SHM), estimating the future loading environments for that system, and then predicting through simulation and past experience the remaining useful life of the system. It is important therefore to distinguish between usage monitoring and health monitoring.

* Author for correspondence (farrar@lanl.gov).

One contribution of 15 to a Theme Issue 'Structural health monitoring'.

© 2006 The Royal Society

- *Usage monitoring*: the process of acquiring operational loading data from a structure or system, which preferably includes a measure of environmental conditions (e.g. temperature and moisture) and operational variables such as mass or speed.
- *Health monitoring*: the process of identifying the presence and quantifying the extent of damage in a system based on information extracted from the measured system response.

In practice both techniques are required for DP, as future life predictions are a function of loads and performance of the structure or system as well as its current condition. Such predictions will be probabilistic in nature. This article summarizes motivations for the development of a DP capability and will discuss the requisite components of a general DP process. This discussion will show how SHM fits into the DP process and concludes by identifying some of the key emerging technologies that will have to be coupled with SHM to obtain a robust DP capability.

2. Motivation for damage prognosis

As with SHM, the interest in DP solutions is based on this technology's tremendous potential for life-safety and/or economic benefits. DP has applications to almost all engineered structures and mechanical systems including those associated with all types of defence hardware, civil infrastructure, manufacturing equipment and commercial aerospace systems. As an example, airframe and jet engine manufacturers are moving to a business model where they lease their hardware to the user, through so-called 'power by the hour' arrangements. Increased profits are then realized by having the ability to assess damage and predict when the damage will reach some critical level that will require corrective action. With such predictions, the owners can better plan their maintenance schedule and optimize the amount of time the hardware is available for leasing, which in turn optimizes the revenue generating potential of these assets. In addition, manufacturers of other large capital expenditure equipment such as earth moving equipment for mining operations would like to move to a business model, whereby they lease the equipment based on the portion of its life that is used rather than on a time-based leasing arrangement. Such a business model requires the ability to monitor the system's response and predict the damage accumulation during a lease interval as a function of the failure damage level.

A distinction that should be made for aircraft is the differing requirements for structure, power plant and systems. For each category, the support schedule specifies maintenance specific items (MSIs) and safety specific items (SSIs). These are specific components that require regular scheduled inspection to avoid either operational disruption or loss of safety. The selection of monitored components is further complicated by adjacent component failures, an example of which is turbine fan-blade-off (FBO). An occurrence of FBO in flight causes an out of balance in the engine; however, despite powering down the engine, the aerodynamic loads cause the fan assembly to 'windmill'. Therefore, the outcome of FBO is forced vibration through the whole airframe, so even though the prognostic sensor system may adequately describe the behaviour of

the engine, the airframe will have absorbed additional fatigue loading without localized monitoring. Although the focus of prognostic methods has been on power plant and structural components as discrete systems, the FBO example highlights the increasing importance of taking a 'multi-component' level approach to DP. However, the sophistication and scope of such an approach currently remains an intractable issue. A further consideration is the treatment of 'systems'—rather than structural and power plant components. Aircraft systems may take the form of data networks, printed circuit boards or recursive (self-learning) software. These examples are typical of the challenges presented for systems prognostics. While the strategy for structural DP has a defined road map, many of the issues facing aircraft systems are still to be resolved. The importance of this area of DP should not be underestimated as for civil aircraft, over the next 10 years, the cost of the system elements will rise to 50% of the purchase cost, while for military platforms the figure will be closer to two-thirds of the total cost.

For civil infrastructure, there is a need for prognosis of structures (e.g. bridges and high-rise buildings) subjected to large-scale discrete events such as earthquakes. As an example, some buildings subjected to the 1995 Kobe earthquake were evaluated for 2 years before reoccupation. Clearly, there is a need to perform more timely and quantified structural condition assessments and then confidently predict how these structures will respond to future loading such as the inevitable aftershocks that occur following a major seismic event. For manufacturing facilities, the current slow post-earthquake assessment and reoccupation process can have an extreme economic impact far beyond the reconstruction costs. This economic impact adversely affects the facility owners as well as companies that insure such facilities.

3. The current state of damage prognosis

The earlier articles in this theme issue show that there is a considerable amount of literature that focuses on identifying damage in engineering systems. The most advanced damage detection systems, which have made the transition from research to practice, include those used for helicopters' gearbox monitoring, where the Federal Aviation Administration has already endorsed their effectiveness and those used to monitor damage accumulation in rotating machinery such as the integrated condition assessment system deployed on US Navy ships (DiUlio *et al.* 2003). However, in almost all cases prognosis remains elusive. To date, one of the few attempts at integrating DP around a predictive capability is also encountered in the field of rotating machinery (Anon. 2000). Successful applications of rotating machinery DP exist because extensive datasets are available, some of which include the monitoring of the machines to failure. Also, for this application the damage location and damage types as well as operational and environmental conditions are often well known *a priori* and do not vary significantly.

Perhaps, the most refined form of combined health and usage monitoring can be found in the helicopter industry. Although it is somewhat short of physical damage-based prognostics, the use of vibration data trending for predictive maintenance can lead to increased rotor component life of 15% (Silverman 2005).

Of even more significance, the introduction of health and usage monitoring systems (HUMS) for main rotor and gearbox components on large rotorcraft has been shown to reduce 'the fatal hull loss within the UK to half what could have otherwise been expected had HUMS not been installed' (McColl 2005). The essential features of this apparent success are that the rotor speed—although not the torque—is maintained typically within 2% of nominal for all flight regimes and that there is a single load path with no redundancy. These constraints provide a basis for a stable vibration spectrum from which a change in measured parameter is attributable to component deterioration. This luxury, consisting of an easily identifiable parametric change coupled with a stable excitation source, makes helicopters particularly amenable to prognostics. Unfortunately, the DP task is complicated when the loading spectrum is constantly varying as would typically be encountered in automotive and military engine applications.

Seismic probabilistic risk assessment (PRA) and seismic margins assessments, as they have been applied to commercial nuclear power plant structures and systems, can to some degree be viewed as forms of DP that has been practiced for more than 20 years. The objective of a seismic PRA is to obtain an estimate of the annual probability that some level of damage will occur to a system as a result of some future earthquake. As an example, seismic PRAs may be used to estimate the probability of core melt in the power plant. Alternatively, the PRAs can be used to estimate the consequences of core melt such as the exposure of the neighbouring population to radioactive materials. One input to a seismic PRA is a probabilistic description of the expected failure rates of system-critical components as a function of earthquake ground motion levels. These probabilistic failure rates are based on analysis, testing or past experience. Additional inputs to the PRA include a probabilistic description of the future earthquake ground motion levels and a fault tree/event tree system model that predicts failure (e.g. core melt) as a result of individual component failures. Historically, the frequency of failure estimate has been used to assist in risk-based maintenance and upgrade of safety class components. For this application, the relative contributions of the various components to system failure are assessed. This information is then used to prioritize equipment for safety upgrades. There has been recent pressure by licensees to move to a risk-based licensing approach and, as a result, the US Nuclear Regulatory Commission has conducted research efforts to establish a risk-based licensing methodology (USNRC 1995). It should be noted that in its current state seismic PRA is carried out strictly as an analytical study without experimental verification on a system level (Ellingwood 1994) and it does not make use of information from any kind of SHM systems. In other words, this process does not account for the inevitable deterioration of the system that will result from normal operations over extended periods of time.

When one looks beyond the examples cited above, few journal papers can be found that discuss DP for other applications. A recently published book (Inman *et al.* 2005), which is based on a series of papers presented at 2003 Pan-American Advanced Study Institute dedicated to the topic of DP is one of the first publications dedicated to the topic of DP. This dearth of publications on DP indicates that this technology is still in the early developmental phase and that there is a need for considerable DP technology development.

4. Defining the damage prognosis problem

The definition of the DP problem starts by first answering three general questions.

(i) What are the loading conditions that cause the damage of concern?
(ii) What techniques should be used to assess and quantify the damage?
(iii) Once the damage has been assessed, what is the goal of the prognosis?

When developing answers to these questions, one will have to consider the length and time-scales associated with the damage and its propagation. As discussed below, the categories do not have sharp boundaries and many applications will overlap the various categories.

For each potential failure mode, the loading conditions that cause damage and subsequent failure fall into three general categories. The first category is gradual wear, where damage accumulates slowly at the material or component level, often on the microscopic scale. An example of such gradual damage accumulation is the corrosion of metallic structural components. The second category is predictable discrete events. While the damage typically still originates on the microscopic scale, it accumulates at faster rates during sudden events that can be characterized *a priori*. Aircraft landings can be viewed as a predictable discrete event that can eventually lead to damage accumulation in the landing gear or airframe. Unpredictable discrete events are the third category in which unknown and severe loading is applied to the system at unpredictable times. Many natural-phenomena hazards such as earthquakes and hurricanes as well as human-made hazards associated with terrorist bomb blasts can produce such unpredictable discrete events.

After identifying the type(s) and source(s) of damage, it is then important to determine which techniques should be used in the damage assessment. The first question that arises concerns whether the assessment should be done online in near real time, or off-line at discrete intervals, as this consideration will strongly influence the data acquisition and data processing requirements, as well as the set limits on the computational requirements of potential assessment and prognosis techniques. For unpredictable discrete events, the assessment must be done online to be of any use, thus limiting the choice of the assessment techniques. However, for gradual wear, there are cases where the assessment need not be performed in near real time, and hence there is much more flexibility to develop an appropriate assessment technique.

Assessment techniques can generally be classified as either physics-based or data-based, though typically a combination of the two will usually be employed. The physics-based assessments are especially useful for predicting system response to new loading conditions and/or new system configurations (damage states). However, physics-based assessments are typically more computationally intensive than data-based techniques.

Data-based assessment techniques, on the other hand, rely on previous measurements from the system to assess the current damage state, typically by means of some sort of pattern recognition method. However, although data-based assessment techniques may be able to indicate a change in the presence of new loading conditions or system configurations, they will perform poorly when trying to classify the nature of the change. Typically, the balance between

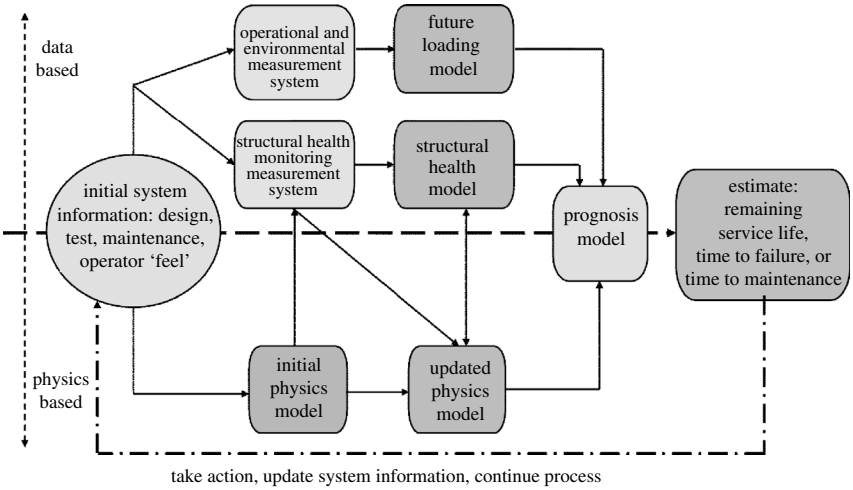

Figure 1. The general components of a DP process.

physics-based models and data-based techniques will depend on the amount of relevant data available and the level of confidence in the predictive accuracy of the physics-based models.

Once the current damage state has been assessed, the prognosis problem can begin to be addressed by determining the goal for the prognosis. Perhaps, the most obvious and desirable type of prognosis estimates how much time remains until maintenance is required, the system fails or the system is no longer usable. Because predictive models typically have more uncertainty associated with them when the structure responds in a nonlinear manner as will often be the case when damage accumulates, *an alternative goal might be to estimate how long the system can continue to safely perform in its anticipated environments before one no longer has confidence in the predictive capabilities of the models that are being used to perform the prognosis.*

5. The damage prognosis process

The general components of a DP process are depicted in figure 1 where the process has been divided into the portions that are physics- and data-based. The DP process begins by collecting as much initial system information as possible including testing and analyses that were performed during the system design as well as maintenance and repair information that might be available. This information is used to develop initial physics-based numerical models of the system as well as to develop the sensing system that will be used for damage assessment and whatever additional sensors are needed to monitor operational and environmental conditions. The physics-based models will also be used to define the necessary sensing system properties (e.g. parameter, location, bandwidth and sensitivity). For instance, an understanding of the physics of gear wear has led to a measure of oil conductivity in helicopter gearboxes on the basis of the metallic contamination after wear and erosion.

Since there will be a finite budget for sensing, the physics-based models will often be a part of an optimization study that will attempt to maximize the observability of damage given constraints on the sensing system properties. As data become available from the sensing systems, they are used to validate and update the physics-based models. These data along with output from the physics-based models will also be used to perform SHM where the existence, location, type and extent of damage are quantified. Data from the operational and environmental sensors are used to develop data-based models that predict the future system loading. The output of the future loading model, SHM model, and the updated physics-based model will all be input into a reliability-based predictive tool that estimates the remaining system life. Note that the definition of 'remaining life' can take on a variety of meanings depending on the specific application. A key point illustrated in figure 1 is that various models will have to be employed in the prognosis process. Moreover, the data- and physics-based portions of the process are not independent. This combination of physics- and data-based models is the key distinguishing attribute of a prognostic rather than a health monitoring system; the capacity to revisit design life prediction on the basis of usage data *and* a change in physical properties. This form of prognostics process allows for alternative load path calculations—arising from redundant systems—and a change in operational use that has not been anticipated in the original design assumptions. Finally, the solution process will be iterative, relying on the assessment of past prediction accuracy to improve future predictions.

6. Emerging technologies that will have an impact on the damage prognosis process

There are many emerging technologies that will have an impact on the development of a DP capability for various types of engineering systems. Those technologies listed below are not intended to be an exhaustive list, but rather illustrative in nature.

(*a*) *Damage measurement systems*

Clearly, some of the most rapidly evolving technologies that will impact the ability to perform DP are associated with sensing, processing and telemetry hardware. There are extensive efforts underway at both academic and corporate research centres to develop large-scale, self-organizing and embedded sensing networks for a wide variety of applications. These studies focus on developing cost-effective dense sensing arrays and novel approaches to powering the sensing systems that harvest the ambient energy available from the structure's operating environment (Sodano *et al.* 2003). Although hardware technologies show every prospect of delivering such systems, their application must be related to the physics of the problem, that is, the damage measurement system must embed sensors that are sensitive to a change in structural condition; and the system itself must be more reliable than the structure/component it is monitoring.

(b) Statistical inference for damage diagnosis

Statistical inference is concerned with the implementation of algorithms that analyse the distribution of extracted features in an effort to make decisions on damage diagnosis and prognosis. The algorithms used in statistical model development fall into the three general categories: (i) group classification, (ii) regression analysis, and (iii) outlier detection. The appropriate algorithm to use will depend on the ability to perform supervised or unsupervised learning. Supervised learning refers to the case where examples of data from damaged and undamaged structures are available. Unsupervised learning refers to the case where data are only available from the undamaged structure. The success of decision-making can be assessed by (i) overall misclassification rate (false positive/negative indications of damage or system failure), (ii) receiver operating characteristic (Egan 1975) curves (ROC), and (iii) confidence intervals on prediction.

One of the main issues in this decision-making procedure is to establish decision threshold values. In particular, extreme value statistics can be employed for the establishment of decision boundaries to minimize false positive and negative indications of damage. Statistical inference is often based on the assumption that the underlying distribution of data is Gaussian. However, the assumption of normality imposes potentially misleading behaviour on the extreme values of the data, namely, those points in the tails of the distribution. As the problem of damage identification specifically focuses attention on these tails, the assumption of normality is likely to lead any analyses astray. The physical interpretation justifying a guarded confidence in the tail distribution is that these values are likely to be influenced by nonlinearities in the structure. A component suffering degradation or approaching failure most likely will have changed from its original design parameters through localized nonlinear deformation. An alternative approach based on extreme value statistics can be applied specifically to model behaviour in the tails of the distribution of interest (Worden *et al.* 2002).

(c) Prediction modelling for future loading estimates

A successful DP requires the measurements of the current system state and the prediction of the system deterioration when subjected to future loading. Based on the analysis of previous loading histories, future loading is forecast using various data-driven time-series prediction modelling techniques. For example, metamodelling such as state-space representation (Ljung 1999) and multivariate ARMA models (Box *et al.* 1994) can be employed to track previous loading and to predict future loading for this purpose. Thus, reliability-based decision analysis provides an appropriate tool to synthesize all this information (Ang & Tang 1984).

(d) Model verification and validation

As DP solutions rely on the deployment of a predictive capability, the credibility of numerical simulations must be established. The process of establishing this confidence in the predictive capabilities of the numerical simulations is accomplished through various activities, collectively referred to as verification and validation (V&V). A significant challenge here is to validate nonlinear models. However, the current state-of-the-art (Hemez *et al.* 2004) in this area is still not at the stage where linear dynamic and stress models are routinely validated, particularly for complex materials such as composites.

(*e*) *Reliability analysis for DP decision-making*

In reliability analysis, the failure state of a system is represented by a function of the response known as the limit state. Then, the probability of failure is the integral of the joint probability density function (JPDF) over the unsafe region bounded by the limit state (Robertson & Hemez 2005). As an example, the objective of a probabilistic reliability analysis may be to answer the question of how many more fatigue cycles the structure can experience before the damage reaches a critical size. If failure is defined, for example, in terms of a wing flutter condition, then the reliability analysis consists of estimating the probability of reaching this limit state (e.g. reduction of the wing's first torsion mode damping), given uncertainties about the model that predicts this frequency reduction as a function of future loading, current health of the system and expected future loading on the wing. Decision-making relies on the estimation of reliability, as well as a quantification of its confidence, to decide which course of action should be taken.

This analysis begins with identifying the failure modes (such as delamination in a composite material) and the random variables that contribute to these failure modes (such as projectile impact velocity, ply orientation angles, homogenized elasticity parameters and material density). To calculate the probability of failure, the JPDF must be integrated across all random variables for the failure region. Because closed form representations of the failure region are generally not available, integration must be approximated by applying Monte Carlo sampling or approximate expansion methods to the previously identified metamodels. Reliability analysis will necessarily be applied to estimate the remaining useful life of the systems under uncertainty.

7. Concluding comments

The challenge of DP is developing and integrating sensing hardware, data interrogation software and predictive modelling software that will prove more robust than the component- and system-level hardware the DP system is intended to identify. This paper aims to provide an overview of the issues that must be addressed and technical approaches being used to realize solutions to this problem. Certainly, considerable technical and cultural challenges remain. The technological aspects are easier to define and anticipate. Already a substantial body of evidence indicates that the individual components of a DP process are realizable, but the integration of all the necessary technologies has been very limited. However, if robust DP solutions are to be adopted, then a change in certification culture must arise that embraces an iterative safety and maintenance process, which may alter significantly from the original design calculations. A 'morphing' safety case based on PRA coupled with a nonlinear iterative model validation will require significant investments in test and analysis correlation studies before proof of robustness is established in each of the constituent technologies. Therefore, it is crucial to apply this technology initially to problems with well-defined damage concerns and where the prognosis system is not relied upon to make decisions impacting life-safety. As an example, deployment of a DP system on an unmanned aerial vehicle may be more appropriate than deployment of such system on a commercial passenger aircraft.

In any case, it is almost certain that new DP approaches will initially be applied in parallel with current system evaluation and maintenance procedures until the DP methodology can be shown to provide a reliable and more cost-effective approach to system operation, assessment and maintenance.

As this technology evolves, it is anticipated that the DP solutions developed through rigorous validation of each technological component will be used to confirm system-level integrity to normal and extreme loading environments; to estimate the probability of mission completion and personnel survivability; to determine the optimal times for preventive maintenance; and to develop the appropriate design or operation modifications that prevent observed damage propagation. The multidisciplinary and challenging nature of the DP problem, its current embryonic state of development, and its tremendous potential for life-safety and economic benefits qualify DP as a grand challenge problem for engineers in the twenty-first century.

References

Ang, A. H. & Tang, W. H. 1984 *Probability concepts in engineering planning and design. Decision, risk, and reliability*, vol. II. New York, NY: Wiley.

Anon. 2000 In *Proc. 13th Int. Congress on Condition Monitoring and Diagnostic Eng. Management, Houston, TX.*

Box, G. E., Jenkins, G. M. & Reinsel, G. C. 1994 *Time series analysis—forecasting and control.* Englewood Cliffs, NJ: Prentice Hall.

DiUlio, M., Savage, C., Finley, B. & Schneider, E. 2003 Taking the integrated condition assessment system to the year 2010. *13th Int. Ship Control Systems Symposium, Orlando, FL.*

Ellingwood, B. 1994 *Validation of seismic probabilistic risk assessment of nuclear power plants*, Division of Engineering, Office of Nuclear Regulatory Research, US Nuclear Regulatory Commission.

Egan, J. P. 1975 *Signal detection theory and ROC analysis.* New York, NY: Academic Press.

Farrar, C. R., *et al.* 2003 Damage prognosis: current status and future needs. Los Alamos National Laboratory report LA-14051-MS.

Hemez, F. M., Doebling, S. W. & Anderson, M. C. 2004 A brief tutorial on verification and validation. In *Proc. 22nd SEM Int. Modal Analysis Conference, Dearborn, MI.*

Inman, D. J., Farrar, C. R., Steffan, V. & Lopes, V. (eds) 2005 *Damage prognosis for aerospace, civil and mechanical systems.* New York, NY: Wiley.

Ljung, L. 1999 *System identification-theory for the user.* Upper Saddle River, NJ: Prentice Hall.

McColl, J. 2005 HUMS in the era of CAA, JAA, EASA and ICAO. AIAC Conference, Melbourne, Paper 112.

Robertson, A. & Hemez, F. 2005 Reliability methods. In *Damage prognosis for aerospace, civil and mechanical systems* (ed. D. J. Inman, C. R. Farrar, V. Steffan, V. Lopes *et al.*), pp. 221–234. New York, NY: Wiley.

Silverman, H., March 2005 T-HUMS—AH64 Lead the fleet (LTF) summary and glimpse at hermes 450 MT-HUMS. AIAC Conference, Melbourne, Paper 151.

Sodano, H. A., Park, G. & Inman, D. J. 2003 A review of power harvesting using piezoelectric materials. *Shock Vibration Digest* **36**, 197–206. (doi:10.1177/0583102404043275)

USNRC 1995 Use of probabilistic risk assessment methods in nuclear regulatory activities; final policy statement. *Federal Register* **60**, 42 622–42 629.

Worden, K., Allen, D. W., Sohn, H. & Farrar, C. F. 2002 Extreme value statistics for damage detection in mechanical structures. Los Alamos National Laboratory Report, LA-13903-MS.