18

18 Health Monitoring and Prognostics

As a result of intense global competition, companies are considering novel approaches to enhance the operational efficiency of their products. For many products and systems, high in-service reliability can be a means to ensure customer satisfaction. In addition, global competitive demands for increased warranties, and the severe liability of product failures, is encouraging manufacturers to improve field reliability and operational availability,[1] provide knowledge of in-service use, and life-cycle operational and environmental conditions.

The American Heritage Dictionary defines prognostic as an adjective that relates to prediction or foretelling and as a noun for a sign or symptom indicating the future course of a disease or sign or forecast of some future occurrence. Hippocrates founded the 21 axioms of prognostics some 2400 years ago.[2] The goal of prognostics is to foretell (predict) the future health (or state) of a system. Health for human beings is defined as a state of complete physical, mental, and social well-being. These ideas can also be applied for the overall health or quality of products and systems. Interest has been growing in monitoring the ongoing health of products and systems in order to provide advance warning failure, and assist in administration and logistics. Here, health is defined as the extent of degradation or deviation from an expected normal condition. Prognostics is the prediction of the future state of health based on current and historical health conditions (Vichare and Pecht 2006). Prognostics deals with prediction of quality in systems. Quality is defined in dictionaries as the essential character or attribute of an entity. It's the inherent characteristic or attribute of something. Thus prognostics deals with prediction of some desired quality or characteristic of a system. Prognostics is based on understanding the science of degradation of the underlying system. This is also called as physics or chemistry or biology or psychology of failure from the viewpoint of the customer. As such, development of sensors and monitoring devices are key for Prognostics and System Health Management (PSHM).

[1]Operational availability is the degree (expressed as a decimal between 0 and 1, or the percentage equivalent) to which a piece of equipment or system can be expected to work properly when required. Operational availability is often calculated by dividing uptime by the sum of uptime and downtime.
[2]See http://classics.mit.edu/Hippocrates/ prognost.html (MIT 2010) (accessed February 2010).

Reliability Engineering, First Edition. Kailash C. Kapur and Michael Pecht.

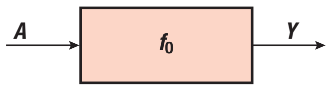

Figure 18.1 Ideal process.

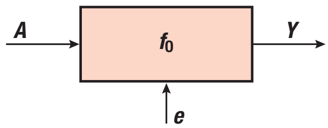

Figure 18.2 Process with variability e (where e represents noise).

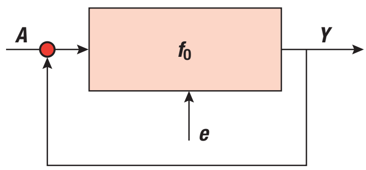

Figure 18.3 Process with feedback loop.

Electronics are integral to the functionality of most systems today, and their reliability is often critical for system reliability (Vichare et al. 2007). This chapter provides a basic understanding of prognostics and health monitoring of products and systems and the techniques being developed to enable prognostics for electronic systems.

18.1 Conceptual Model for Prognostics

Figure 18.1 shows a typical system with inputs A [single or vector] and response variable (or output) represented by Y. If we have perfect knowledge about this system, and we know the transfer function $f_0(A) = Y$, then inputs A can be determined as $A = f_0^{-1}(Y)$. If we know the system (the transfer function), then we can predict the response variable Y and adjust the inputs A to maintain the output within the desired range. This is the ideal deterministic process and is shown in Figure 18.1.

A perfect prognostics is the situation where we know the transfer function and we have perfect knowledge of the system. In that case we can foretell many measures of Y when A is the system input. If we know what output Y is desired, we can determine how input A should be adjusted. There are challenges to achieve this goal:

(a) The inverse problem is not unique and not easy to determine.
(b) We often lack knowledge (or there is uncertainty) about our model.
(c) The real-world systems might be very complex and cause output Y to appear as a random variable.

A cause of variation in Y is due to error or noise factors represented by e in Figure 18.2. Thus, e makes the output Y a random variable. In terms of reliability, Y may be the time to failure random variable.

Thus, if the output is not closer to the ideal value or the target due to the presence of noise factors, we can measure it and one way to overcome these deviations is to use feedback as shown in Figure 18.3. It is clear that feedback is reactive and can be too late in order to adjust the input properly. Instead, we want to be proactive and

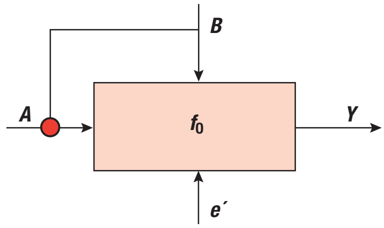

Figure 18.4 Prognostics P with feed forward.

prevent deviations from ideal or target value occurring in the response Y in the first place.

The traditional approaches have been based on feedback to correct the system behavior. Many of the reliability methodologies are based on feedback, like reliability growth through testing.

In prognostics, we need to understand the underlying causes of error. We decompose error e further to identify the disturbance or another factor B that is part of e and that affect the output. Thus, now we have

$$e = (e', B). \tag{18.1}$$

We can measure this disturbance B (though we cannot change or control disturbance) so we can determine how to change the other input variables to create feed forward and maintain the system response variable Y closer to the target.

We can track trends and make forecasts to identify future system behavior when disturbances are measured. The trend is to use artificial neural networks to model prognostics. Based on the system knowledge, we can use feed forward to provide prognostics. We account for error (uncertainties, perturbations, and disturbances) using feedback, and we use the known disturbance and its measurement for feed forward and prognostics. Figure 18.4 illustrates the role of prognostic system P and its feed forward loop B.

The design of prognostics P, relies on feed forward model to properly regulate inputs A. The role of prognostic system P is to develop the relationship $B = g(X_1, X_2, X_3, \ldots, X_n)$ such that the response variable $Y = f_0(A, B)$ remains in the expected or desired range.

The design of prognostic systems is challenging. Finding the inverse function is not always easy and not necessarily unique. The actual system and its environment in the future can be quite different from what we perceive today. In addition, there is a lack of knowledge about the inner workings of the system in operation. This lack of knowledge stems from uncertainty and could be caused by the initial model error or not considering all factors. As a result the response variable Y appears as a random variable to the observer.

In Figure 18.5, another illustration for a feed forward design is shown. The feed forward prognostics model incorporates the concept of disturbance B and error e'. We monitor and measure disturbance B and want to know its effect on the output. Likewise, we are interested in understanding the reasons for error e'; for example, whether it is due to variation or uncertainty that stems from incomplete or fuzzy understanding of the system's operating and environmental factors. Our goal is to decompose the causes of error into informative variables with the aid of the subject matter experts. Finally, we design the prognostic system such that by monitoring B,

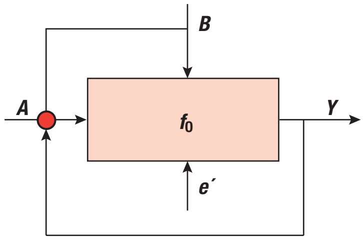

Figure 18.5 Process with feedback and feed forward.

when values of *B* change, we adjust the input variables *A* accordingly in order to prevent a problem or to prolong system lifetime. This is illustrated in Figure 18.5, where both feedback loop and feed forward are used to maintain the response *Y* in the expected range.

Identifying prognostics information is based on considering a system's use information in several areas: in environmental and operational conditions, measured process values, monitored residuals, and fault alarms to identify fault modes. We have to keep in mind that different fault modes can progress to manifest into different failures and hence require different prognostic models.

Among many methods of prognostics is estimating the remaining useful life (RUL) of a system. RUL can be estimated from historical and operational data collected from a system. Various methods are used to determine system degradation and predict RUL.

Another method is estimating the probability of failure (POF). POF is the failure probability distribution of the system or a component. Additionally, we can study time to failure (TTF), the time a component is expected to fail. TTF defines the time when a system no longer meets its design specifications.

Prognostic methods combine RUL, TTF, and POF with other techniques to extend system life, ensure mission completion, and improve corporate profitability.

18.2 Reliability and Prognostics

Reliability is the ability of a product or system to perform as intended (i.e., without failure and within specified performance limits) for a specified time, in its life-cycle environment. Traditional reliability prediction methods for electronic products include Mil-HDBK-217 (U.S. Department of Defense 1965), 217-PLUS, Telcordia (Telcordia Technologies 2001), PRISM (Denson 1999), and FIDES (FIDES Group 2004). These methods rely on the collection of failure data and generally assume the components of the system have failure rates (most often assumed to be constant) that can be modified by independent "modifiers" to account for various quality, operating, and environmental conditions. There are numerous well-documented concerns with this type of modeling approach (Cushing et al. 1993; Leonard 1991b; Talmor and Arueti 1997; Wong 1990). The general consensus is that these handbooks should never be used, because they are inaccurate for predicting actual field failures and provide highly misleading predictions, which can result in poor designs and logistics decisions (Morris 1990; Wong 1990).

The traditional handbook method for the reliability prediction of electronics started with Mil-Hdbk-217A, published in 1965. In this handbook, there was only a single-point failure rate for all monolithic integrated circuits, regardless of the stresses, the materials, or the architecture. Mil-Hdbk-217B was published in 1973, with the RCA/Boeing models simplified by the U.S. Air Force to follow a statistical exponential (constant failure rate) distribution. Since then, all the updates were mostly "band-aids" for a modeling approach that was proven to be flawed (Pecht and Nash 1994). In 1987–1990, the Center for Advanced Life Cycle Engineering (CALCE) at the University of Maryland was awarded a contract to update Mil-Hdbk-217. It was concluded that this handbook should be cancelled and the use of this type of modeling approach discouraged.

In 1998, IEEE 1413 standard, "IEEE Standard Methodology for Reliability Prediction and Assessment for Electronic Systems and Equipment," was approved to provide guidance on the appropriate elements of a reliability prediction (IEEE Standard 1413–1998 1998). A companion guidebook, IEEE 1413.1, "IEEE Guide for Selecting and Using Reliability Predictions Based on IEEE 1413," provides information and an assessment of the common methods of reliability prediction for a given application (IEEE Standard 1413.1-2002 2003). It is shown that the Mil-Hdbk-217 is flawed. There is also discussion of the advantage of reliability prediction methods that use stress and damage physics-of-failure (PoF) technique.

The PoF approach and design-for-reliability (DfR) methods have been developed by CALCE (Pecht and Dasgupta 1995) with the support of industry, government, and other universities. PoF is an approach that utilizes knowledge of a product's life-cycle loading and failure mechanisms to perform reliability modeling, design, and assessment. The approach is based on the identification of potential failure modes, failure mechanisms, and failure sites for the product as a function of its life-cycle loading conditions. The stress at each failure site is obtained as a function of both the loading conditions and the product geometry and material properties. Damage models are then used to determine fault generation and propagation.

Prognostics and health management (PHM) is a method that permits the assessment of the reliability of a product (or system) under its actual application conditions. When combined with physics-of-failure models, it is thus possible to make continuously updated predictions based on the actual environmental and operational conditions. PHM techniques combine sensing, recording, interpretation of environmental, operational, and performance-related parameters to indicate a system's health. PHM can be implemented through the use of various techniques to sense and interpret the parameters indicative of:

- Performance degradation, such as deviation of operating parameters from their expected values
- Physical or electrical degradation, such as material cracking, corrosion, interfacial delamination, or increases in electrical resistance or threshold voltage
- Changes in a life-cycle profile, such as usage duration and frequency, ambient temperature and humidity, vibration, and shock.

The framework for prognostics is shown in Figure 18.6. Performance data from various levels of an electronic product or system can be monitored in situ and analyzed using prognostic algorithms. Different implementation approaches can be adopted individually or in combination. These approaches will be discussed in the subsequent sections. Ultimately, the objective is to predict the advent of failure in

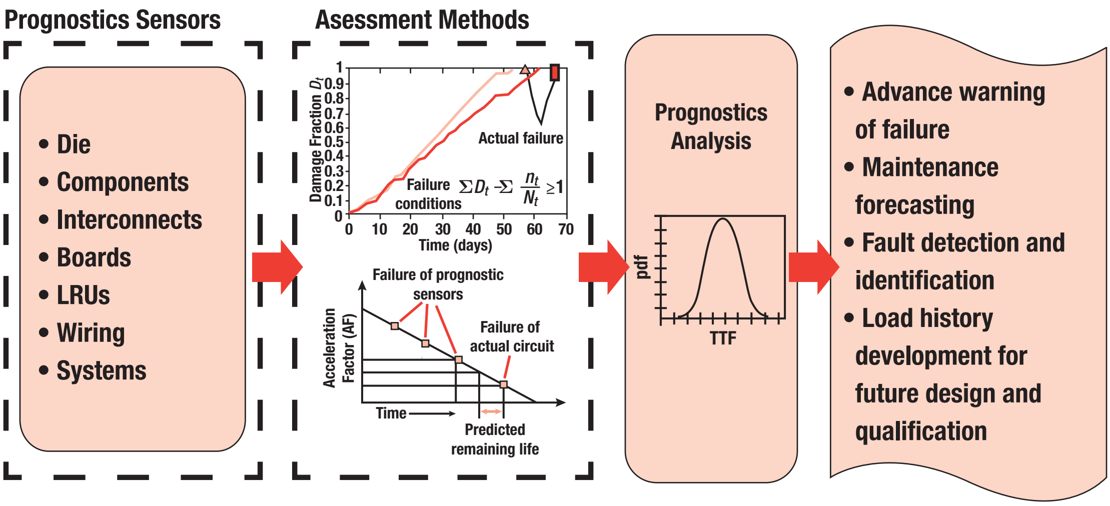

Figure 18.6 Framework for prognostics and health management.

terms of a distribution of remaining life, level of degradation, or probability of mission survival.

18.3 PHM for Electronics

Most products and systems contain significant electronics content to provide needed functionality and performance. If one can assess the extent of deviation or degradation from an expected normal operating condition for electronics, this information can be used to meet several powerful goals, which include:

1. Providing advanced warning of failures;
2. Minimizing unscheduled maintenance, extending maintenance cycles, and maintaining effectiveness through timely repair actions;
3. Reducing the life-cycle cost of equipment by decreasing inspection costs, downtime, and inventory; and
4. Improving qualification and assisting in the design and logistical support of fielded and future systems (Vichare and Pecht 2006).

In other words, since electronics are playing an increasingly large role in providing operational capabilities for today's products and systems, prognostic techniques have become highly desirable.

Some of first efforts in diagnostic health monitoring of electronics involved the use of built-in test (BIT), defined as an onboard hardware–software diagnostic means to identify and locate faults. A BIT can consist of error detection and correction circuits, totally self-checking circuits, and self-verification circuits (Vichare and Pecht 2006). Two types of BIT concepts are employed in electronic systems: interruptive BIT

(I-BIT) and continuous BIT (C-BIT). The concept behind I-BIT is that normal equipment operation is suspended during BIT operation. The concept behind C-BIT is that equipment is monitored continuously and automatically without affecting normal operation.

Several studies (Johnson 1996; Pecht et al. 2001) conducted on the use of BIT for fault identification and diagnostics showed that BIT can be prone to false alarms and can result in unnecessary costly replacement, requalification, delayed shipping, and loss of system availability. BIT concepts are still being developed to reduce the occurrence of spurious failure indications. However, there is also reason to believe that many of the failures actually occurred, but were intermittent in nature (DoD 5000.2 Policy Document 2004). The persistence of such issues over the years is perhaps because the use of BIT has been restricted to low volume systems. Thus, BIT has generally not been designed to provide prognostics or remaining useful life due to accumulated damage or progression of faults. Rather, it has served primarily as a diagnostic tool.

PHM has also emerged as one of the key enablers for achieving efficient system-level maintenance and lowering life-cycle costs in military systems. In November 2002, the U.S. Deputy Under Secretary of Defense for Logistics and Materiel Readiness released a policy called condition-based maintenance plus (CBM+). CBM+ represents an effort to shift unscheduled corrective equipment maintenance of new and legacy systems to preventive and predictive approaches that schedule maintenance based upon the evidence of need. A 2005 survey of 11 CBM programs highlighted "electronics prognostics" as one of the most needed maintenance-related features or applications, without regard for cost (Cutter and Thompson 2005), a view also shared by the avionics industry (Kirkland et al. 2004). Department of Defense 5000.2 policy document on defense acquisition, which states that "program managers shall optimize operational readiness through affordable, integrated, embedded diagnostics and prognostics, embedded training and testing, serialized item management, automatic identification technology, and iterative technology refreshment" (DoD 5000.2 Policy Document 2004). Thus, a prognostics capability has become a requirement for any system sold to the U.S. Department of Defense.

Prognostics and health management is also emerging as a high priority issue in space applications. NASA's Ames Research Center (ARC) in California is focused on conducting fundamental research in the field of Integrated Systems Health Management (ISHM). ARC is involved in design of health management systems, selection and optimization of sensors, in situ monitoring, data analysis, prognostics, and diagnostics. The prognostics center for excellence at ARC develops algorithms to predict the remaining life of NASA's systems and subsystems. ARC's current prognostics projects involve power semiconductor devices (investigation of the effects of ageing on power semiconductor components, identification of failure precursors to build a PoF model and development of algorithms for end-of-life prediction), batteries(algorithms for batteries prognosis), flight actuators (PoF modeling and development of algorithms for estimation of remaining life), solid rocket motor failure prediction, and aircraft wiring health management (Korsmeyer 2013).

In addition to in-service reliability assessment and maintenance, health monitoring can also be effectively used to support product take-back and end-of-life decisions. Product take-back indicates the responsibility of manufacturers for their products over the entire life cycle, including disposal. The motivation driving product take-back is the concept of Extended Producer Responsibility (EPR) for postconsumer electronic waste (Rose et al. 1999). The objective of EPR is to make manufacturers and distributors financially responsible for their products when they are no longer needed.

End-of-life product recovery strategies include repair, refurbishing, remanufacturing, reuse of components, material recycling, and disposal. One of the challenges in end-of-life decision-making is to determine whether product lines can be extruded, whether any components could be reused and what subset should be disposed of in order to minimize system costs (Sandborn and Murphy 1999). Several interdependent issues must be considered concurrently to properly determine the optimum component reuse ratio, including assembly/disassembly costs and any defects introduced by either process, product degradation incurred in the original life cycle, and the waste stream associated with the life cycle. Among these factors, the estimate of the degradation of the product in its original life cycle could be the most uncertain input to end-of-life decisions. This could be effectively carried out using health monitoring, with knowledge of the entire history of the product's life cycle.

Scheidt and Zong (1994) proposed the development of special electrical ports, referred to as green ports, to retrieve product usage data that could assist in the recycling and reuse of electronic products. Klausner et al. (1998a, 1998b) proposed the use of an integrated electronic data log (EDL) for recording parameters indicative of product degradation. The EDL was implemented on electric motors to increase the reuse of motors. In another study (Simon et al. 2000), domestic appliances were monitored for collecting usage data by means of electronic units fitted on the appliances. This work introduced the Life Cycle Data Acquisition Unit, which can be used for data collection and also for diagnostics and servicing. Middendorf et al. (2002) suggested developing life-information modules to record the cycle conditions of products for reliability assessment, product refurbishing, and reuse.

Designers often establish the usable life of products and warranties based on extrapolating accelerated test results to assumed usage rates and life-cycle conditions. These assumptions may be based on worst-case scenarios of various parameters composing the end user environment. Thus, if the assumed conditions and actual use conditions are the same, the product would last for the designed time, as shown in Figure 18.7(a). However, this is rarely true, and usage and environmental conditions could vary significantly from those assumed. For example, consider products equipped with life-consumption monitoring systems for providing in situ assessment of remaining life. In this situation, even if the product is used at a higher usage rate and in harsh conditions, it can still avoid unscheduled maintenance and catastrophic failure, maintain safety, and ultimately save cost. These are typically the motivational factors for use of health monitoring or life consumption monitoring (LCM), as shown in Figure 18.7(b).

One of the vital inputs in making end-of-life decisions is the estimate of degradation and the remaining life of the product. Figure 18.7c illustrates a scenario in which a working product is returned at the end of its designed life. Using the health monitors installed within the product, the reusable life can be assessed. Unlike testing conducted after the product is returned, this estimate can be made without having to disassemble the product. Ultimately, depending on other factors, such as cost of the product, demand for spares, cost, and yield in assembly and disassembly, the manufacturer can choose to reuse or dispose.

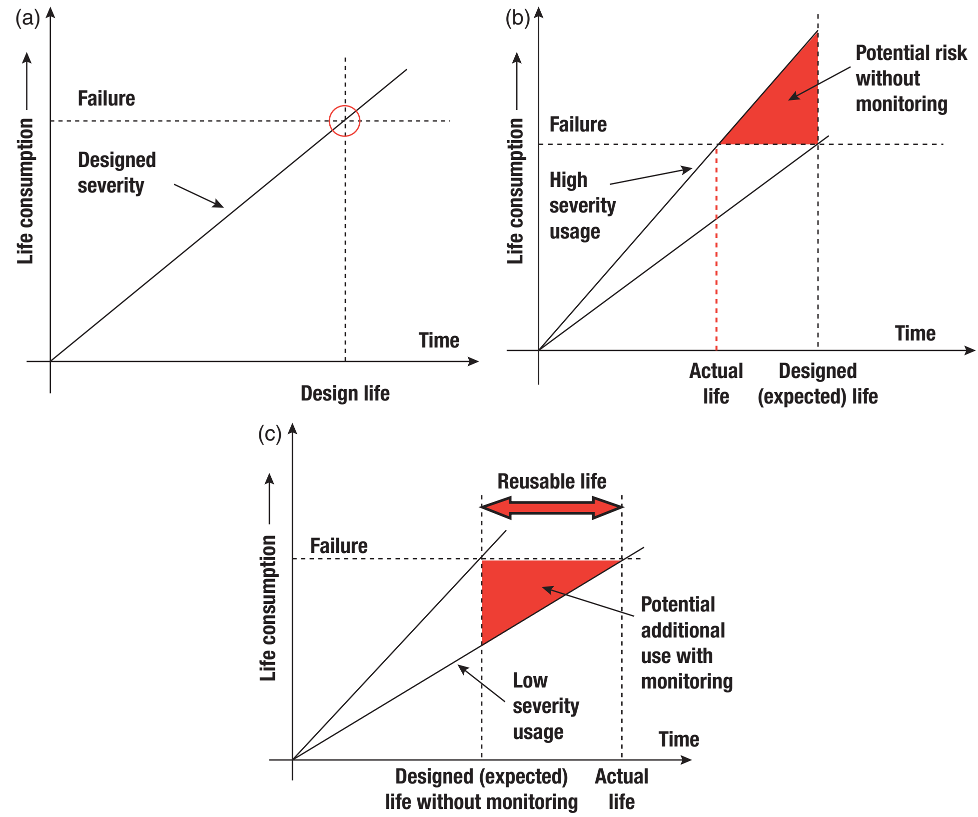

Figure 18.7 Application of health monitoring for product reuse.

18.4 PHM Concepts and Methods

The general PHM methodology is shown in Figure 18.8 (Gu and Pecht 2007). The first step involves a virtual life assessment, where design data, expected life-cycle conditions, failure modes, mechanisms, and effects analysis (FMMEA), and PoF models are the inputs to obtain a reliability (virtual life) assessment. Based on the virtual life assessment, it is possible to prioritize the critical failure modes and failure mechanisms. The existing sensor data, bus monitor data, maintenance, and inspection record can also be used to identify the abnormal conditions and parameters. Based on this information, the monitoring parameters and sensor locations for PHM can be determined.

Based on the collected operational and environmental data, the health status of the products can be assessed. Damage can also be calculated from the PoF models to obtain the remaining life. Then PHM information can be used for maintenance forecasting and decisions that minimize life-cycle costs, maximize availability, or some other utility function.

The different approaches to prognostics and the state of research in electronics PHM are presented here. Three current approaches include:

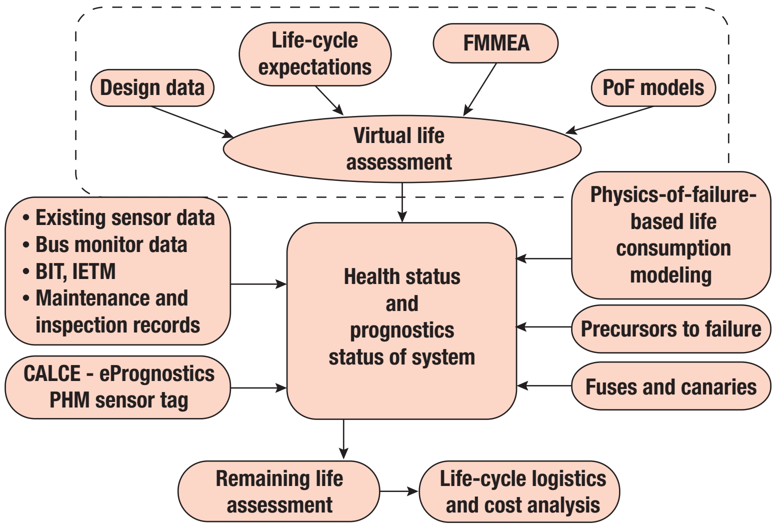

Figure 18.8 PHM methodology.

1. The use of fuses and canary devices
2. Monitoring and reasoning of failure precursors
3. Monitoring environmental and usage condition for stress and damage PoF modeling.

18.4.1 Fuses and Canaries

Expendable devices, such as fuses and canaries, have been a traditional method of protection for structures and electrical power systems. Fuses and circuit breakers are examples of elements used in electronic products to sense excessive current drain and to disconnect power. Fuses within circuits safeguard parts against voltage transients or excessive power dissipation, and protect power supplies from shorted parts. For example, thermostats can be used to sense critical temperature limiting conditions, and to shut down the product, or a part of the system, until the temperature returns to normal. In some products, self-checking circuitry can also be incorporated to sense abnormal conditions and to make adjustments to restore normal conditions, or to activate switching means to compensate for a malfunction (Ramakrishnan et al. 2000).

The word "canary" is derived from one of coal mining's earliest systems for warning of the presence of hazardous gas using the canary bird. Because the canary is more sensitive to hazardous gases than humans, the death or sickening of the canary was an indication to the miners to get out of the shaft. The canary thus provided an effective early warning of catastrophic failure that was easy to interpret. The same approach has been employed in prognostic health monitoring. Canary devices mounted on the actual product can also be used to provide advance warning of failure due to specific wearout failure mechanisms.

Mishra and Pecht (2002) studied the applicability of semiconductor-level health monitors by using precalibrated cells (circuits) located on the same chip with the actual circuitry. The prognostics cell approach, known as Sentinel Semiconductor™ technology, has been commercialized to provide an early warning sentinel for upcoming device failures (Ridgetop Semiconductor-Sentinel Silicon™ Library 2004). The prognostic cells are available for 0.35-, 0.25-, and 0.18-μm CMOS processes; the power consumption is approximately 600 microwatts. The cell size is typically 800 μm^2 at the 0.25-μm process size. Currently, prognostic cells are available for semiconductor failure mechanisms, such as electrostatic discharge (ESD), hot carrier, metal migration, dielectric breakdown, and radiation effects.

The time to failure of prognostic canaries can be precalibrated with respect to the time to failure of the actual product. Because of their location, these canaries contain and experience substantially similar dependencies, as does the actual product. The stresses that contribute to degradation of the circuit include voltage, current, temperature, humidity, and radiation. Since the operational stresses are the same, the damage rate is expected to be the same for both the circuits. However, the prognostic canary is designed to fail faster through increased stress on the canary structure by means of scaling.

Scaling can be achieved by controlled increase of the stress (e.g., current density) inside the canaries. With the same amount of current passing through both circuits, if the cross-sectional area of the current-carrying paths in the canary is decreased, a higher current density is achieved. Further control in current density can be achieved by increasing the voltage level applied to the canaries. A combination of both of these techniques can also be used. Higher current density leads to higher internal (joule) heating, causing greater stress on the canaries. When a current of higher density passes through the canaries, they are expected to fail faster than the actual circuit (Mishra and Pecht 2002).

Figure 18.9 shows the failure distribution of the actual product and the canary health monitors. Under the same environmental and operational loading conditions, the canary health monitors wear out faster to indicate the impending failure of the actual product. Canaries can be calibrated to provide sufficient advance warning of failure (prognostic distance) to enable appropriate maintenance and replacement activities. This point can be adjusted to some other early indication level. Multiple trigger points can also be provided, using multiple canaries spaced over the bathtub curve.

Goodman et al. (2006) used a prognostic canary to monitor time-dependent dielectric breakdown (TDDB) of the metal-oxide semiconductor field-effect transistor (MOSFET) on the integrated circuits. The prognostic canary was accelerated to

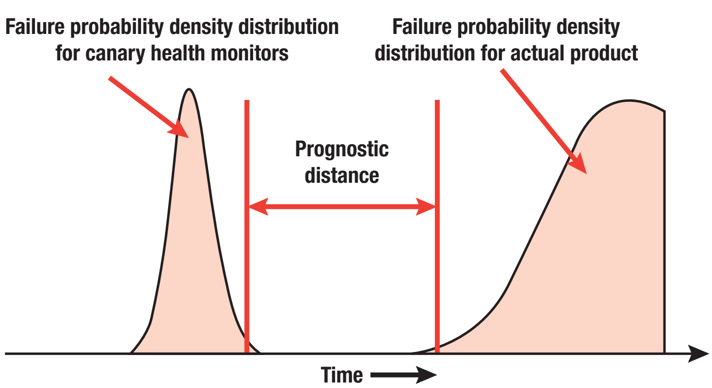

Figure 18.9 Advanced warning of failure using canary structures.

failure under certain environmental conditions. Acceleration of the breakdown of an oxide could be achieved by applying a voltage higher than the supply voltage, to increase the electric field across the oxide. When the prognostics canary failed, a certain fraction of the circuit lifetime was used up. The fraction of consumed circuit life was dependent on the amount of overvoltage applied and could be estimated from the known distribution of failure times.

The extension of this approach to board-level failures was proposed by Anderson and Wilcoxon (2004), who created canary components (located on the same printed circuit board) that include the same mechanisms that lead to failure in actual components. Anderson et al. identified two prospective failure mechanisms: (1) low cycle fatigue of solder joints, assessed by monitoring solder joints on and within the canary package; and (2) corrosion monitoring, using circuits that are susceptible to corrosion. The environmental degradation of these canaries was assessed using accelerated testing, and degradation levels were calibrated and correlated to actual failure levels of the main system. The corrosion test device included an electrical circuitry susceptible to various corrosion-induced mechanisms. Impedance spectroscopy was proposed for identifying changes in the circuits by measuring the magnitude and phase angle of impedance as a function of frequency. The change in impedance characteristics can be correlated to indicate specific degradation mechanisms.

There remain unanswered questions with the use of fuses and canaries for PHM. For example, if a canary monitoring a circuit is replaced, what is the impact when the product is reenergized? What protective architectures are appropriate for postrepair operations? What maintenance guidance must be documented and followed when fail-safe protective architectures have or have not been included? The canary approach is also difficult to implement in legacy systems, because it may require requalification of the entire system with the canary module. Also, the integration of fuses and canaries with the host electronic system could be an issue with respect to real estate on semiconductors and boards. Finally, the company must ensure that the additional cost of implementing PHM can be recovered through increased operational and maintenance efficiencies.

18.5 Monitoring and Reasoning of Failure Precursors

A failure precursor is a data event or trend that signifies impending failure. A precursor indication is usually a change in a measurable variable that can be associated with subsequent failure. For example, a shift in the output voltage of a power supply might suggest impending failure due to a damaged feedback regulator and opto-isolator circuitry. Failures can then be predicted by using causal relationships between measured variables that can be correlated with subsequent failure, and for PoF.

A first step in failure precursor PHM is to select the life-cycle parameters to be monitored. Parameters can be identified based on factors that are crucial for safety, that are likely to cause catastrophic failures, that are essential for mission completeness, or that can result in long downtimes. Selection can also be based on knowledge of the critical parameters established by past experience, field failure data on similar products, and on qualification testing. More systematic methods, such as FMMEA (Ganesan et al. 2005b), can also be used to determine parameters that need to be monitored.

Table 18.1 Potential failure precursors for electronics

Electronic subsystem	Failure precursor
Switching power supply	■ DC output (voltage and current levels) ■ Ripple ■ Pulse width duty cycle ■ Efficiency ■ Feedback (voltage and current levels) ■ Leakage current ■ RF noise
Cables and connectors	■ Impedance changes ■ Physical damage ■ High-energy dielectric breakdown
CMOS IC	■ Supply leakage current ■ Supply current variation ■ Operating signature ■ Current noise ■ Logic level variations
Voltage-controlled oscillator	■ Output frequency ■ Power loss ■ Efficiency ■ Phase distortion ■ Noise
Field effect transistor	■ Gate leakage current/resistance ■ Drain-source leakage current/resistance
Ceramic chip capacitor	■ Leakage current/resistance ■ Dissipation factor ■ RF noise
General purpose diode	■ Reverse leakage current ■ Forward voltage drop ■ Thermal resistance ■ Power dissipation ■ RF noise
Electrolytic capacitor	■ Leakage current/resistance ■ Dissipation factor ■ RF noise
RF power amplifier	■ Voltage standing wave ratio (VSWR) ■ Power dissipation ■ Leakage current

Pecht et al. (1999) proposed several measurable parameters that can be used as failure precursors for electronic products, including switching power supplies, cables and connectors, CMOS integrated circuits, and voltage-controlled high-frequency oscillators (see Table 18.1).

In general, to implement a precursor reasoning-based PHM system, it is necessary to identify the precursor variables for monitoring, and then develop a reasoning algorithm to correlate the change in the precursor variable with the impending failure. This characterization is typically performed by measuring the precursor variable under an expected or accelerated usage profile. Based on the characterization, a model is developed—typically a parametric curve-fit, neural network, Bayesian network, or a time-series trending of a precursor signal. This approach assumes that there is one or more expected usage profiles that are predictable and can be simulated, often in a laboratory setup. In some products, the usage profiles are predictable, but this is not always true.

For a fielded product with highly varying usage profiles, an unexpected change in the usage profile could result in a different (noncharacterized) change in the precursor signal. If the precursor reasoning model is not characterized to factor in the uncertainty in life-cycle usage and environmental profiles, it may provide false alarms. Additionally, it may not always be possible to characterize the precursor signals under all possible usage scenarios (assuming they are known and can be simulated). Thus, the characterization and model development process can often be time-consuming and costly and may not always work.

There are many examples of the monitoring and trending of failure precursor to assess health and product reliability. Some key studies are presented in the next section.

Smith and Campbell (2000) developed a quiescent current monitor (QCM) that can detect elevated Iddq current in real time during operation.[3] The QCM performed leakage current measurements on every transition of the system clock to get maximum coverage of the IC in real time. Pecuh et al. (1999) and Xue and Walker (2004) proposed a low-power built-in current monitor for CMOS devices. In the Pecuh et al. study, the current monitor was developed and tested on a series of inverters for simulating open and short faults. Both fault types were successfully detected and operational speeds of up to 100 MHz were achieved with negligible effect on the performance of the circuit under test. The current sensor developed by Xue and Walker enabled Iddq monitoring at a resolution level of 10 pA. The system translated the current level into a digital signal with scan chain readout. This concept was verified by fabrication on a test chip.

GMA Industries (Wright and Kirkland 2003; Wright et al. 2001, 2003) proposed embedding molecular test equipment (MTE) within ICs to enable them to continuously test themselves during normal operation and to provide a visual indication that they have failed. The molecular test equipment could be fabricated and embedded within the individual integrated circuit in the chip substrate. The molecular-sized sensor "sea of needles" could be used to measure voltage, current, and other electrical parameters, as well as sense changes in the chemical structure of integrated circuits that are indicative of pending or actual circuit failure. This research focuses on the development of specialized doping techniques for carbon nanotubes to form the basic structure comprising the sensors. The integration of these sensors within conventional IC circuit devices, as well as the use of molecular wires for the interconnection of sensor networks, is an important factor in this research. However, no product or prototype has been developed to date.

[3]The power supply current (Idd) can be defined by two elements: the Iddq-quiescent current and the Iddt-transient or dynamic current. Iddq is the leakage current drawn by the CMOS circuit when it is in a stable (quiescent) state. Iddt is the supply current produced by circuits under test during a transition period after the input has been applied. Iddq has been reported to have the potential for detecting defects such as bridging, opens, and parasitic transistor defects. Operational and environmental stresses, such as temperature, voltage, and radiation, can quickly degrade previously undetected faults and increase the leakage current (Iddq). There is extensive literature on Iddq testing, but little has been done on using Iddq for in situ PHM. Monitoring Iddq has been more popular than monitoring Iddt.

Kanniche and Mamat-Ibrahim (2004) developed an algorithm for health monitoring of voltage source inverters with pulse width modulation. The algorithm was designed to detect and identify transistor open circuit faults and intermittent misfiring faults occurring in electronic drives. The mathematical foundations of the algorithm were based on discrete wavelet transform (DWT) and fuzzy logic (FL). Current waveforms were monitored and continuously analyzed using DWT to identify faults that may occur due to constant stress, voltage swings, rapid speed variations, frequent stop/start-ups, and constant overloads. After fault detection, "if-then" fuzzy rules were used for VLSI fault diagnosis to pinpoint the fault device. The algorithm was demonstrated to detect certain intermittent faults under laboratory experimental conditions.

Self-monitoring analysis and reporting technology (SMART), currently employed in select computing equipment for hard disk drives (HDD), is another example of precursor monitoring (Hughes et al. 2002; Self-Monitoring Analysis and Reporting Technology (SMART) 2001). HDD operating parameters, including the flying height of the head, error counts, variations in spin time, temperature, and data transfer rates, are monitored to provide advance warning of failures (see Table 18.2). This is achieved through an interface between the computer's start-up program (BIOS) and the hard disk drive.

Systems for early fault detection and failure prediction are being developed using variables such as current, voltage, and temperature, continuously monitored at various locations inside the system. Sun Microsystems refers to this approach as continuous system telemetry harnesses. Along with sensor information, soft performance parameters, such as loads, throughputs, queue lengths, and bit error rates, are tracked. Prior to PHM implementation, characterization is conducted by monitoring the signals of different variables to establish a multivariate state estimation technique (MSET) model of the "healthy" systems. Once the "healthy" model is established using this

Table 18.2 Monitoring parameters based on reliability concerns in hard drives

Reliability issues	Parameters monitored

■ Head assembly
  ■ Crack on head
  ■ Head contamination or resonance
  ■ Bad connection to electronics module
■ Motors/bearings
  ■ Motor failure
  ■ Worn bearing
  ■ Excessive run-out
  ■ No spin
■ Electronic module
  ■ Circuit/chip failure
  ■ Interconnection/solder joint failure
  ■ Bad connection to drive or bus
■ Media
  ■ Scratch/defects
  ■ Retries
  ■ Bad servo
■ ECC corrections
	■ Head flying height: A downward trend in flying height will often precede a head crash.
	■ Error checking and correction (ECC) use and error counts: The number of errors encountered by the drive, even if corrected internally, often signals problems developing with the drive.
	■ Spin-up time: Changes in spin-up time can reflect problems with the spindle motor.
	■ Temperature: Increases in drive temperature often signal spindle motor problems.
	■ Data throughput: Reduction in the transfer rate of data can signal various internal problems.

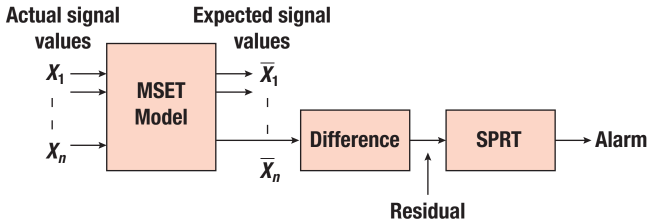

Figure 18.10 Sun Microsystems' approach to PHM.

data, it is used to predict the signal of a particular variable based on learned correlations among all variables (Whisnant et al. 2005). Based on the expected variability in the value of a particular variable during application, a sequential probability ratio test (SPRT) is constructed. During actual monitoring, SPRT is used to detect deviations of the actual signal from the expected signal based on distributions (and not on a single threshold value) (Cassidy et al. 2002; Mishra and Gross 2003). This signal is generated in real time based on learned correlations during characterization (Figure 18.10). A new signal of residuals is generated, which is the arithmetic difference of the actual and expected time-series signal values. These differences are used as input to the SPRT model, which continuously analyzes the deviations and provides an alarm if the deviations are of concern (Whisnant et al. 2005). The monitored data is analyzed to provide alarms based on leading indicators of failure, and enable use of monitored signals for fault diagnosis, root-cause analysis, and analysis of faults due to software aging (Vaidyanathan and Gross 2003).

Brown et al. (2005) demonstrated that the remaining useful life of a commercial global positioning system (GPS) can be predicted by using a precursor-to-failure approach. The failure modes for GPS included precision failure due to an increase in position error, and solution failure due to increased outage probability. These failure progressions were monitored in situ by recording system-level features reported using the National Marine Electronics Association (NMEA) Protocol 0183. The GPS was characterized to collect the principal feature value for a range of operating conditions. Based on experimental results, parametric models were developed to correlate the offset in the principal feature value with solution failure. During the experiment, the BIT provided no indication of an impending solution failure (Brown et al. 2005).

18.5.1 Monitoring Environmental and Usage Profiles for Damage Modeling

The life-cycle profile of a product consists of manufacturing, storage, handling, operating, and nonoperating conditions. The life-cycle loads (Table 18.3), either individually or in various combinations, may lead to performance or physical degradation of the product and reduce its service life (Ramakrishnan and Pecht 2003). The extent and rate of product degradation depends upon the magnitude and duration of exposure (usage rate, frequency, and severity) to such loads. If one can measure these loads in situ, the load profiles can be used in conjunction with damage models to assess the degradation due to cumulative load exposures.

The assessment of the impact of life-cycle usage and environmental loads on electronic structures and components was studied by Ramakrishnan and Pecht (2003). This study introduced the LCM methodology (Figure 18.11), which combined in situ

Table 18.3 Examples of life-cycle loads

Load	Load conditions
Thermal	Steady-state temperature, temperature ranges, temperature cycles, temperature gradients, ramp rates, heat dissipation
Mechanical	Pressure magnitude, pressure gradient, vibration, shock load, acoustic level, strain, stress
Chemical	Aggressive versus inert environment, humidity level, contamination, ozone, pollution, fuel spills
Physical	Radiation, electromagnetic interference, altitude
Electrical	Current, voltage, power, resistance

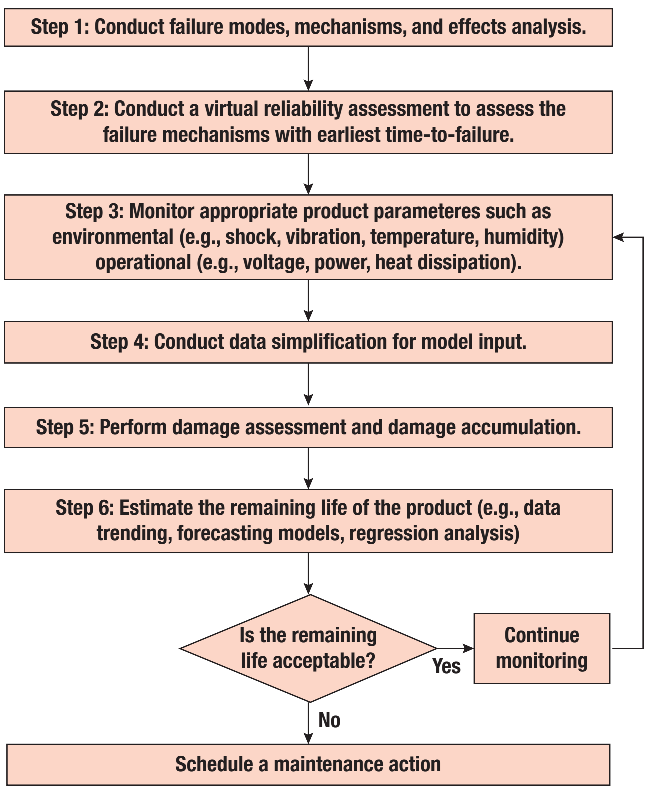

Figure 18.11 CALCE life consumption monitoring methodology.

measured loads with physics-based stress and damage models to assess remaining product life.

Mathew et al. (2006) applied the LCM methodology to conduct a prognostic remaining-life assessment of circuit cards inside a space shuttle solid rocket booster (SRB). Vibration-time history, recorded on the SRB from the prelaunch stage to splashdown, was used in conjunction with physics-based models to assess damage. Using the entire life-cycle loading profile of the SRBs, the remaining life of the components and structures on the circuit cards were predicted. It was determined that an electrical failure was not expected within another 40 missions. However, vibration and shock analysis exposed an unexpected failure of the circuit card due to a broken aluminum bracket mounted on the circuit card. Damage accumulation analysis determined that the aluminum brackets had lost significant life due to shock loading.

Shetty et al. (2002) applied the LCM methodology to conduct a prognostic remaining-life assessment of the end effector electronics unit (EEEU) inside the robotic arm of the space shuttle remote manipulator system (SMRS). A life-cycle loading profile of thermal and vibrational loads was developed for the EEEU boards. Damage assessment was conducted using physics-based mechanical and thermomechanical damage models. A prognostic estimate using a combination of damage models, inspection, and accelerated testing showed that there was little degradation in the electronics and they could be expected to last another 20 years.

Gu et al. (2007) developed a methodology for monitoring, recording, and analyzing the life-cycle vibration loads for remaining-life prognostics of electronics. The responses of printed circuit boards to vibration loading in terms of bending curvature were monitored using strain gauges. The interconnect strain values were then calculated from the measured PCB response and used in a vibration failure fatigue model for damage assessment. Damage estimates were accumulated using Miner's rule after every mission and then used to predict the life consumed and remaining life. The methodology was demonstrated for remaining-life prognostics of a printed circuit board assembly. The results were also verified by checking the resistance data.

In case studies (Mishra et al. 2002; Ramakrishnan and Pecht 2003), an electronic component-board assembly was placed under the hood of an automobile and subjected to normal driving conditions. Temperature and vibrations were measured in situ in the application environment. Using the monitored environmental data, stress and damage models were developed and used to estimate consumed life. Figure 18.12 shows estimates obtained using similarity analysis, and the actual measured life. Only LCM accounted for this unforeseen event because the operating environment was being monitored in situ.

Vichare and Pecht (2006) outlined generic strategies for in situ load monitoring, including selecting appropriate parameters to monitor and designing an effective monitoring plan. Methods for processing the raw sensor data during in situ monitoring to reduce the memory requirements and power consumption of the monitoring device were presented. Approaches were also presented for embedding intelligent front-end data processing capabilities in monitoring systems to enable data reduction

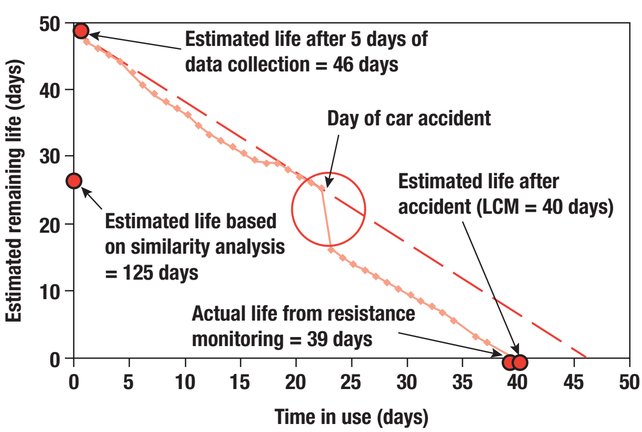

Figure 18.12 Remaining-life estimation of test board.

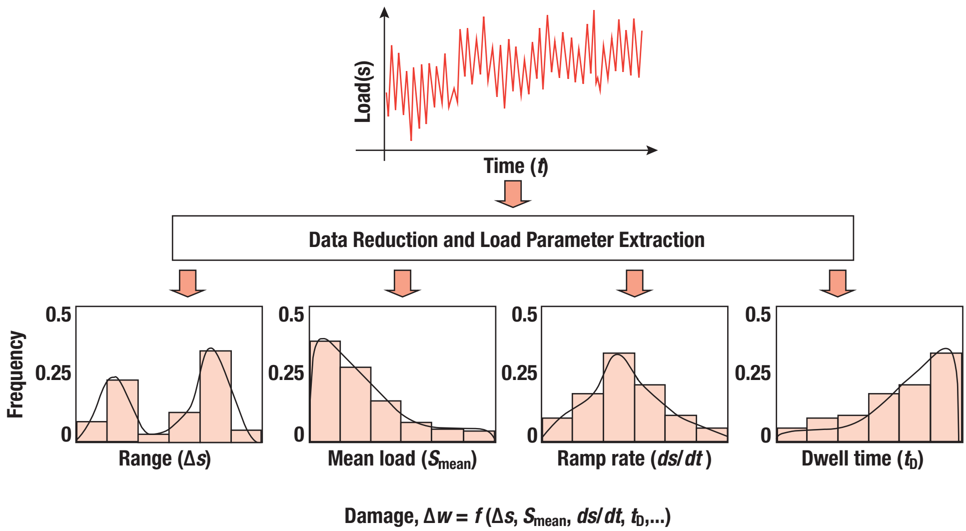

Figure 18.13 Load feature extraction.

and simplification (without sacrificing relevant load information) prior to input in damage models for health assessment and prognostics.

To reduce on-board storage space, power consumption, and uninterrupted data collection over longer durations, Vichare et al. (2006) suggested embedding data reduction and load parameter extraction algorithms into sensor modules. As shown in Figure 18.13, a time-load signal can be monitored in situ using sensors, and further processed to extract cyclic range(s), cyclic mean load (S_{mean}), and rate of change of load (ds/dt), using embedded load extraction algorithms. The extracted load parameters can be stored in appropriately binned histograms to achieve further data reduction. After the binned data are downloaded, it can be used to estimate the distributions of the load parameters. The usage history is used for damage accumulation and remaining life prediction.

Efforts to monitor life-cycle load data on avionics modules can be found in time-stress measurement device (TSMD) studies. Over the years, TSMD designs have been upgraded using advanced sensors, and miniaturized TSMDs are being developed with advances in microprocessor and nonvolatile memory technologies (Rouet and Foucher 2004).

Searls et al. (2001) undertook in situ temperature measurements in both notebook and desktop computers used in different parts of the world. In terms of the commercial applications of this approach, IBM has installed temperature sensors on hard drives (Drive-TIP) to mitigate risks due to severe temperature conditions, such as thermal tilt of the disk stack and actuator arm, off-track writing, data corruptions on adjacent cylinders, and outgassing of lubricants on the spindle motor. The sensor is controlled using a dedicated algorithm to generate errors and control fan speeds.

Strategies for efficient in situ health monitoring of notebook computers were provided by Vichare et al. (2004). In this study, the authors monitored and statistically analyzed the temperatures inside a notebook computer, including those experienced during usage, storage, and transportation, and discussed the need to collect such data both to improve the thermal design of the product and to monitor prognostic health. The temperature data was processed using ordered overall range (OOR) to convert an irregular time–temperature history into peaks and valleys and also to remove noise due to small cycles and sensor variations. A three-parameter Rainflow algorithm was then used to process the OOR results to extract full and half cycles with cyclic range, mean, and ramp rates. The effects of power cycles, usage history, CPU computing resources usage, and external thermal environment on peak transient thermal loads were characterized.

In 2001, the European Union funded a 4-year project, "Environmental Life-Cycle Information Management and Acquisition" (ELIMA), which aimed to develop ways to manage the life cycles of products (Bodenhoefer 2004). The objective of this work was to predict the remaining life time of parts removed from products, based on dynamic data, such as operation time, temperature, and power consumption. As a case study, the member companies monitored the application conditions of a game console and a household refrigerator. The work concluded that in general, it was essential to consider the environments associated with all life intervals of the equipment. These included not only the operational and maintenance environments, but also the preoperational environments, when stresses maybe imposed on the parts during manufacturing, assembly, inspection, testing, shipping, and installation. Such stresses are often overlooked, but can have a significant impact on the eventual reliability of equipment.

Skormin et al. (2002) developed a data-mining model for failure prognostics of avionics units. The model provided a means of clustering data on parameters measured during operation, such as vibration, temperature, power supply, functional overload, and air pressure. These parameters are monitored in situ on the flight using time–stress measurement devices. Unlike the physics-based assessments made by Ramakrishnan (Ramakrishnan and Pecht 2003), the data-mining model relies on statistical data of exposures to environmental factors and operational conditions.

Tuchband and Pecht (2007) presented the use of prognostics for a military line replaceable units (LRU) based on their life-cycle loads. The study was part of an effort funded by the Office of Secretary of Defense to develop an interactive supply chain system for the U.S. military. The objective was to integrate prognostics, wireless communication, and databases through a web portal to enable cost-effective maintenance and replacement of electronics. The study showed that prognostics-based maintenance scheduling could be implemented into military electronic systems. The approach involves an integration of embedded sensors on the LRU, wireless communication for data transmission, a PoF-based algorithm for data simplification and damage estimation, and a method for uploading this information to the Internet. Finally, the use of prognostics for electronic military systems enabled failure avoidance, high availability, and reduction of life-cycle costs.

The PoF models can be used to calculate the remaining useful life, but it is necessary to identify the uncertainties in the prognostic approach and assess the impact of these uncertainties on the remaining life distribution in order to make risk-informed decisions. With uncertainty analysis, a prediction can be expressed as a failure probability.

Gu et al. (2007) implemented the uncertainty analysis of prognostics for electronics under vibration loading. Gu identified the uncertainty sources and categorized them into four different types: measurement uncertainty, parameter uncertainty, failure criteria uncertainty, and future usage uncertainty (see Figure 18.14). Gu et al. (2007)

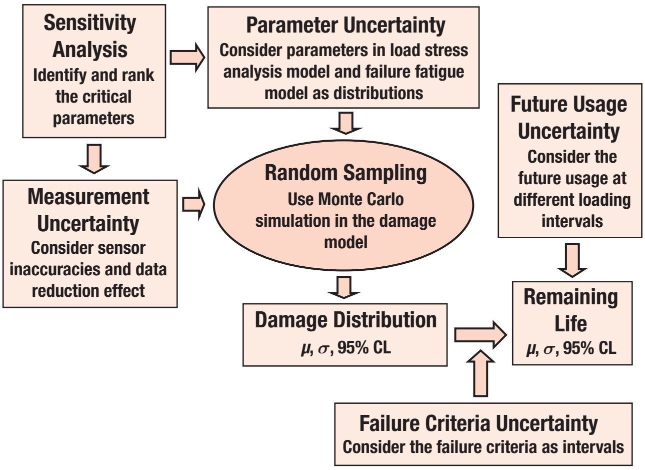

Figure 18.14 Uncertainty implementation for prognostics.

utilized a sensitivity analysis to identify the dominant input variables that influence the model output. With information of input parameter variable distributions, a Monte Carlo simulation was used to provide a distribution of accumulated damage. From the accumulated damage distributions, the remaining life was then predicted with confidence intervals. A case study was also presented for an electronic board under vibration loading and a step-by-step demonstration of the uncertainty analysis implementation. The results showed that the experimentally measured failure time was within the bounds of the uncertainty analysis prediction.

18.6 Implementation of PHM in a System of Systems

System of systems is the term used to describe a complex system comprising of many different subsystems that may be structurally or functionally connected. These different subsystems might themselves be made up of different subsystems. In a system of systems, many independent subsystems are integrated such that the individual functions of the subsystems are combined to achieve a capability/function beyond the capability of the individual subsystems. For example, a military aircraft is made up of subsystems, including airframe, body, engines, landing gear, wheels, weapons, radar, and avionics. Avionic subsystems could include the communication navigation and identification (CNI) system, global positioning system (GPS), inertial navigation system (INS), identification friend or foe (IFF) system, landing aids, and voice and data communication systems.

Implementing an effective PHM strategy for a complete system of systems requires integrating different prognostic and health monitoring approaches. Because the systems are so complex, the first step in implementation of prognostics is to determine the weak link(s) in the system. One of the ways to achieve this is by conducting a FMMEA for the product. Once the FMMEA have been identified, a combination of canaries, precursor reasoning, and life-cycle damage modeling may be implemented for different subsystems of the product, depending on their failure attributes. Once the monitoring techniques have been decided, then the next step is to analyze the data.

Different data analysis approaches, such as data-driven models, PoF-based models, or hybrid data analysis models, can be used to analyze the same recorded data. For example, operational loads of computer system electronics, such as temperature, voltage, current, and acceleration, can be used with PoF damage models to calculate the susceptibility to electromigration between metallization, and thermal fatigue of interconnects, plated-through holes, and die attach. Also, data about the CPU usage, current, and CPU temperature, and so on, can be used to build a statistical model that is based on the correlations between these parameters. This data-driven model can be appropriately trained to detect thermal anomalies and identify signs for certain transistor degradation.

Implementation of prognostics for system of systems is complicated and in the very initial stages of research and development. But there has been tremendous development in the certain areas related to prognostics and health management. Advances in sensors, microprocessors, compact nonvolatile memory, battery technologies, and wireless telemetry have already enabled the implementation of sensor modules and autonomous data loggers. Integrated, miniaturized, low power, reliable sensor systems operated using portable power supplies (such as batteries) are being developed. These sensor systems have a self-contained architecture requiring minimum or no intrusion into the host product, in addition to specialized sensors for monitoring localized parameters. Sensors with embedded algorithms will enable fault detection, diagnostics, and remaining life prognostics, which will ultimately drive the supply chain. The prognostic information will be linked via wireless communications to relay needs to maintenance officers. Automatic identification techniques such as radio frequency identification (RFID) will be used to locate parts in the supply chain, all integrated through a secure web portal to acquire and deliver replacement parts quickly on an as-needed basis.

Research is being conducted in the field of algorithm development to analyze, trend and isolate large-scale multivariate data. Methods, such as projection pursuit using principal component analysis and support vector machines, mahanalobis distance analysis, symbolic time series analysis, neural networks analysis, and Bayesian networks analysis, can be used to process multivariate data.

Even though there are advances in certain areas related to prognostics, many challenges still remain. The key issues with regard to implementing PHM for a system of systems include decisions of which systems within the system of systems to monitor, which system parameters to monitor, selection of sensors to monitor parameters, power supply for sensors, on board memory for storage of sensed data, in situ data acquisition, and feature extraction from the collected data. It is also a challenge to understand how failures in one system affect another system within the system of systems and how it affects the functioning of the overall system of systems. Getting information from one system to the other could be hard especially when the systems are made by different vendors. Other issues that should be considered before implementation of PHM for system of systems are the economic impact due to such a program, contribution of PHM implementation to a condition-based maintenance, and logistics.

The elements necessary for a PHM application are available, but the integration of these components to achieve the prognostics for a system of systems is still in the works. In the future, electronic system designs will integrate sensing and processing modules that will enable in situ PHM. A combination of different PHM implementations for different subsystems of a system of systems will be the norm for the industry.

18.7 Summary

Due to the increasing amount of electronics in the world and the competitive drive toward more reliable products, prognostics and health management is being looked upon as a cost-effective solution for the reliability prediction of electronic products and systems. Approaches for implementing prognostics and health management in products and systems include (1) installing built-in structures (fuses and canaries) that will fail faster than the actual product when subjected to application conditions, (2) monitoring and reasoning of parameters (e.g., system characteristics, defects, and performance) that are indicative of an impending failure, and (3) monitoring and modeling environmental and usage data that influence the system's health and converting the measured data into life consumed. A combination of these approaches may be necessary to successfully assess the degradation of a product or system in real time and subsequently provide estimates of remaining useful life.

Problems

18.1 One of the potential investment returns (cost avoidances) listed for PHM was associated with warranties. Explain the ways in which the cost of warranties could be decreased by using PHM.

18.2 Why does unscheduled maintenance cost more than scheduled maintenance?

18.3 What is "remaining useful life"? How can remaining useful life prognosis improve system reliability?

18.4 What is a failure precursor and how can a failure precursor be identified? Explain with examples.

18.5 Explain the methods for PHM.

18.6 Discuss the pros and cons for data-driven prognostic methods and PoF prognostic methods.

18.7 Suppose you are designing a PHM system for batteries. Discuss the steps and factors for the implementation of the PHM system.