PROGNOSTICS FOR POLYMER POSITIVE TEMPERATURE COEFFICIENT RESETTABLE FUSES

Shunfeng Cheng[1], Kwok Tom[2], and Michael Pecht[1,3*]

[1]Center for Advanced Life Cycle Engineering (CALCE)
University of Maryland, College Park, Maryland, 20742
[2] Army Research Laboratory
2800 Powder Mill Road, Adelphi, Maryland, 20783
[3]Prognostics and Health Management Center
City University of Hong Kong
*Corresponding author

Abstract: As a circuit protection device, polymer positive temperature coefficient (PPTC) resettable fuses have been widely used in over-current or over-temperature circuit-protection designs in computers, automotive circuits, telecommunications equipment, and medical devices. Failure of a resettable fuse can cause abnormal operation of and damage to a circuit. Prognostics for a resettable fuse enables advance warning of the failure of the fuse and the prediction of its remaining useful life; therefore, the operator can take action to maintain or replace the component to reduce damage to the circuit. This paper uses the cross-validation equential probability ratio test (CV-SPRT) method to monitor the failure precursor parameters in situ, and uses an autoregressive integrated moving average (ARIMA) model to predict failure based on the precursors.

Keywords: Polymer positive temperature coefficient (PPTC); Resettable fuse; Prognostics; Failure precursors; Cross-validation combined sequential probability ratio test (CV-SPRT); Autoregressive integrated moving average (ARIMA).

Introduction

A resettable fuse is a circuit protection device. Resettable fuses are widely used in automotive circuits (e.g., the protection of micro-motors in window lifts, seats, and door locks), computers (e.g., the protection of the circuits of hard disk drives, interface ports, USB protection, and cooling fan motors), telecommunication devices (e.g., cell phones), battery packs, power supplies, medical electronics, and so on [1]. The failures or abnormal behaviors of resettable fuses may cause damage to circuits, the abnormal operation of circuits (inability to work at normal current), or unnecessary operations forcing operators to switch off and on the power to reset the circuit.

It is necessary to incorporate prognostics and health management (PHM) to monitor the health status of a fuse to detect abnormal behavior and predict the fuse's remaining useful life (RUL). PHM is an enabling discipline consisting of technologies and methods to assess the reliability of a product in its actual life cycle conditions to determine the advent of failure and mitigate system risk [2][3]. PHM generally combines sensing and interpretation of environmental, operational, and performance-related parameters to assess the health of a product and predict remaining useful life. Assessing the health of a product provides information that can be used to meet several critical goals: (1) providing advance 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 [2][3].

Figure 1 shows the operational process of polymer positive temperature coefficient (PPTC) resettable fuses [4]. Under normal conditions (normal ambient temperature and normal current), a PPTC fuse operates in a low resistance state when the normal current (less than the hold current, which is the maximum steady-state current the fuse can carry without tripping at the ambient temperature) passes through it. When a fault current (higher than the trip current, which is the minimum current that causes a fuse to trip at the ambient temperature) occurs, the resistance of the fuse increases sharply. Because of the sharp increase in resistance, the fuse will decrease the current in order to protect the circuit. The sharp resistance increase is called a trip. After the trip, a PPTC fuse does not break as does a traditional fuse. Instead, it keeps the high resistance state and allows a small trickle current to pass through the circuit. The fuse will start to reset to a low resistance state when the heat or fault current is removed and/or the power is switched off [1][4][5]. Trip time is the time required for a PPTC fuse to decrease the current of the circuit to the hold current at the ambient temperature [1][4][5].

Trip time has been identified as a failure precursor parameter by trip cycle tests [4]. Figure 2 shows the trip time change of sample A with trip cycles at a typical temperature. The PHM for PPTC resettable fuses includes anomaly detection and RUL prediction, both of which can be conducted by applying data-driven methods to the failure precursor parameters. In our study, anomaly detection for a PPTC resettable fuse was conducted by using the sequential probability ratio test (SPRT) on the trip time. However, the traditional SPRT has difficulty selecting model parameters [6][7][8][9], which are critical parameters determining the performance of SPRT. The cross-validation (CV) technique is proposed to select the proper model parameters for SPRT; this is called CV-SPRT. RUL prediction can be implemented by using autoregressive integrated moving average (ARIMA), which is a model for linear time series prediction [10]. The trip time signal in the future is predicted by ARIMA, and RUL can be estimated when the prediction crosses the failure criteria. The failure criterion in terms of trip time is defined based on the specific application. In our current project, the failure of the fuse in terms of trip time is defined as when the trip time decreases or increases by 15% of the mean trip time from 2000 to 4500 cycles. For sample A, the failure criterion is less than 7.5 s or higher than 10.2 s.

Anomaly Detection for PPTC Resettable Fuses

Anomaly detection using CV-SPRT is illustrated in Figure 3. The features of the monitored product in its healthy status are extracted and learned in the training procedure. The training data can be obtained from historic data or from the stable operational phases when the product is known to be healthy. Statistical features of the training data, such as the distribution, are extracted to create a detection baseline. In this procedure, proper model parameters for SPRT are selected by a cross-validation process. In the detection procedure, in-situ monitored data is compared with the baseline sequentially using CV-SPRT to detect the anomalies.

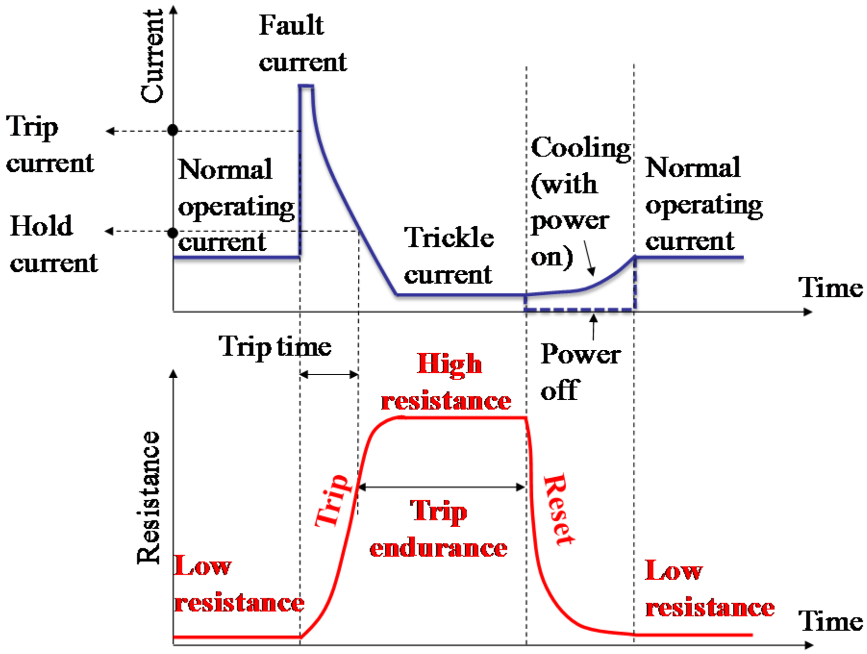

Figure 1: Operational process of resettable fuses.

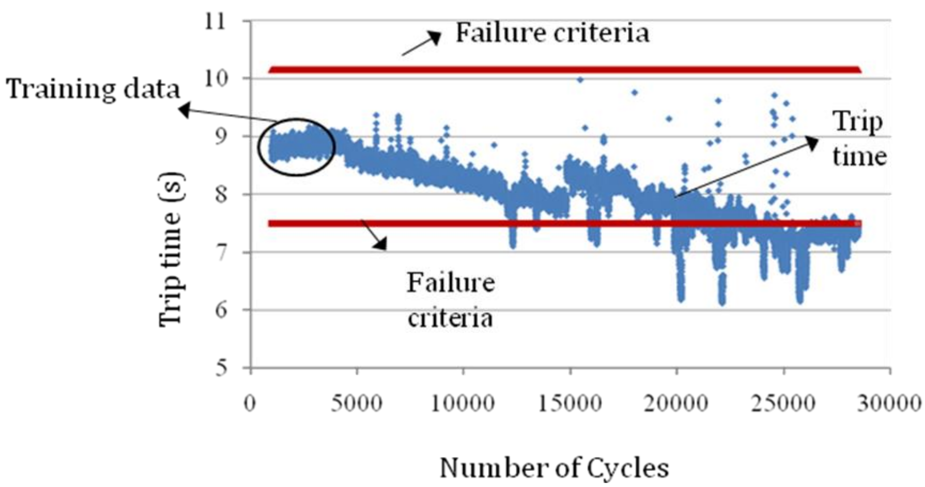

Figure 2: Trip time signals of a fuse (Sample A).

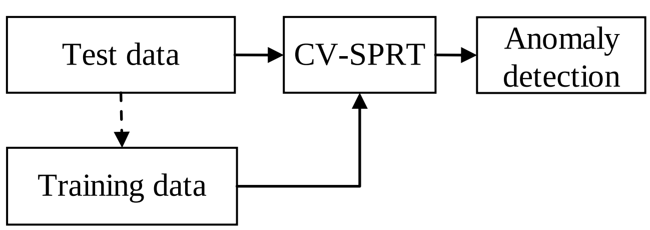

Figure 3: Anomaly detection procedure using CV-SPRT.

Figure 4 shows the procedure of SPRT. The SPRT calculates the SPRT index for each in-situ monitored data point and compares the SPRT index with the boundary to make a decision. The SPRT index is the natural logarithm of the ratio of the probability that accepts the null hypothesis to the probability that accepts the alternative hypothesis. The SPRT index can be calculated as long as the distribution of the detected data is available. For normal distribution, the null (healthy) hypothesis H_0 represents the healthy state, with mean = 0 and standard deviation = σ; the alternative (degraded) hypothesis includes four cases: 1) H_1: the mean of the test data has shifted high to $+M$, with no change in standard deviation; 2) H_2: the mean of the test data has shifted low to $-M$, with no change in standard deviation; 3) H_3: the variance of test data has increased to $V\sigma^2$, with no change in mean; 4) H_4: the variance of test data has decreased to σ^2/V, with no change in mean [11]. M and V are the predetermined system disturbance magnitudes, which are decided by the user, and in general they are several times the standard deviation of the training data. Four SPRT index formulas for a normal distribution are shown as Equations (1) to (4) [11][13].

$$SPRT_1 = \frac{M}{\sigma^2}\sum_{i=1}^{n}(x_i - \frac{M}{2}) \tag{1}$$

$$SPRT_2 = \frac{M}{\sigma^2}\sum_{i=1}^{n}(-x_i - \frac{M}{2}) \tag{2}$$

$$SPRT_3 = \frac{\sum_{i-1}^{n} {x_i}^2}{2\sigma^2}(1-\frac{1}{V}) - (n/2)\ln V \tag{3}$$

$$SPRT_4 = \frac{\sum_{i-1}^{n} {x_i}^2}{2\sigma^2}(1-V) + (n/2)\ln V \tag{4}$$

SPRT uses the probabilities of missed alarms and false alarms to create the thresholds of acceptance and rejection of the null hypothesis, as shown in equation (5). False alarm probability, α , is defined as the probability that H_0 is rejected even though it is true, while missed alarm probability, β , is defined as the probability that H_0 is accepted when it is actually false. Under certain false alarm probabilities and missed alarm probabilities, SPRT gives a decision with minimum sampling.

$$A = \ln(\frac{\beta}{1-\alpha}),\ B = \ln(\frac{1-\beta}{\alpha}) \tag{5}$$

If all SPRT indices are less than the lower limit, the healthy hypothesis (H_0) is accepted, the corresponding indices are reset, and sampling continues. If any SPRT index is larger than the upper limit, the alarm is given, and the corresponding index is reset, and sampling continues. If all SPRT indices are in the range, the information is not sufficient to make a conclusion, and the sampling continues. The comparison of the four SPRT indices with limits is parallel. This procedure ensures that an alarm will be generated when any of the four SPRT indices reaches the upper limit [6].

Four model parameters should be determined for SPRT detection: the system disturbance magnitude, M, which is m times the standard deviation of the training data; the variation factor, V; the false alarm probability, α; and the missed alarm probability, β. No systematic method to select the model parameters is reported in literature. In practice, these model parameters are selected by experience. In this paper, the model parameter set *(m, V, α, β)* for SPRT can be selected by 10-fold CV without the need for experience. Each parameter in the set has a recommended range [12][13], as shown in Table 1.

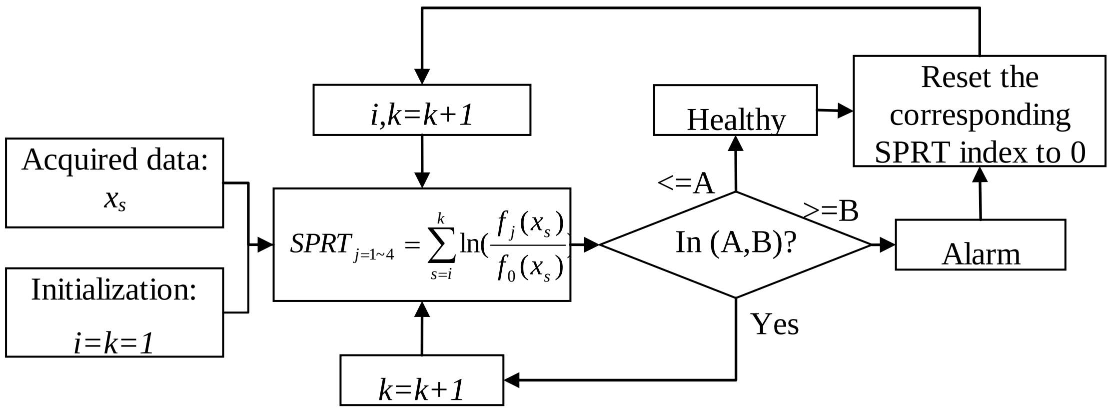

Figure 4: SPRT procedure.

Table 1: Range and Change Interval of Four Parameters

Parameter	Range	Start Value	Interval
m	2 ~ 4	2	1
V	2 ~ 4	2	1
α	0.005 ~ 0.2	0.005	0.005
β	0.005 ~ 0.2	0.005	0.005

The procedure of model parameter selection by 10-fold CV is shown in Figure 5. The first step is to select the labeled dataset, which are all healthy data, as well as the training data for SPRT after the proper model parameter set is obtained. After the labeled data is selected, a model parameter set (m, V, α, β) is assigned to SPRT. For each assigned model parameter set, the actual false alarm probability (α') and missed alarm probability (β') are then calculated by 10-fold CV.

In 10-fold CV, the total labeled data are all healthy data and are partitioned into 10 subsets of equal or nearly equal size. Ten iterations should be conducted to calculate the actual false alarm rate (α') and missed alarm rate (β'). In each of the 10 iterations, one different subset is selected as the original validation data, and the remaining nine subsets are the training data. If SPRT is only run on the original validation data, which are all healthy data, only false alarms can be obtained if SPRT detects some anomalies. In order to calculate the actual missed alarm probability of SPRT, a known abnormal dataset should be added into the original validation dataset to form an updated validation dataset. The abnormal dataset can be generated by randomly picking up data out of the range of [μ-$m\times\sigma$, μ+$m\times\sigma$]. Here the μ and σ are the mean and standard deviation of the training data (the remaining nine subsets of the labeled data).

In each iteration, both false alarms and missed alarms can be identified based on comparison of the detection results with the updated validation data. The false alarm rate and missed alarm rate can be calculated by equations (6) and (7). The process is repeated 10 times until each of the subsets has been selected as the original validation dataset. The mean of the 10 false alarm rates and the mean of the 10 missed alarm rates are calculated by equations (8) and (9) as the final false alarm probability, (α'), and final missed alarm probability, (β'), when using the total labeled data (without the abnormal data) as training data.

The actual false alarm probability and missed alarm probability are then compared with the ones in the selected model parameter set. If both the final false alarm rate and the missed alarm rate are less than the user-specified false alarm probability and missed alarm probability, respectively, the model parameter set should be accepted as a proper model parameter set for SPRT. If not, another model parameter set based on the interval in Table 1 should be re-assigned and the performance of the SPRT with the model parameter should be evaluated by 10-fold CV again. In practice, selection can start from lower values and stop when the acceptance criterion is satisfied. One can also run the 10-fold CV for all the model parameter sets and then select the one with the lowest false and missed alarm probability. A model parameter set should be updated when the labeled (training) data are updated, since the actual false alarm and missed alarm rates are calculated based on the specific labeled (training) data.

$$\alpha_i' = \frac{Number\ of\ false\ alarms\ in\ i^{th}\ iteration}{Number\ of\ updated\ validation\ data\ in\ i^{th}\ iteration} \tag{6}$$

$$\beta_i' = \frac{Number\ of\ missed\ alarms\ in\ i^{th}\ iteration}{Number\ of\ updated\ validation\ data\ in\ i^{th}\ iteration} \tag{7}$$

$$\alpha' = \frac{1}{10}\sum \alpha_i' \tag{8}$$

$$\beta' = \frac{1}{10}\sum \beta_i' \tag{9}$$

The trip times from 2000 to 4500 cycles were selected as the training data for CV-SPRT. The reason the first 2000 cycles were not added was that the fuse was not running stably. Before using CV-SPRT to detect the anomaly, the normality of the training data was checked to determine if the trip time was normally distributed. An abnormal dataset with 50 abnormal data points was randomly generated out of the range of [μ-$m\times\sigma$, μ+$m\times\sigma$]. Here the μ and σ are the mean and standard deviation of the training data (the remaining nine subsets of the labeled data). The abnormal data was added into the original validation data set to form an updated validation data set, which was used to calculate the actual false alarm and missed alarm rates in each iteration. After the 10-fold CV process, the parameter set (m=4, V=2, α =0.01, and β=0.005) was selected (the actual false alarm was estimated at 0.005 and the actual missed alarm was estimated at 0.001). SPRT then used this model parameter set and the total labeled data (without the abnormal data) as training data to detect the anomalies in the test dataset. Figure 6 shows the final detection results of the trip time signals from 2000 to 28000 cycles. Continuous alarms were generated from 5000 cycles. Most of the alarms were trigged by a decrease in the mean trip time and an increase in the variance. The first actual intermittent failure occurred around the 12,000th cycle, and permanent failure occurred around the 20,000th cycle. Thus, SPRT is sensitive enough to detect the anomalies in advance.

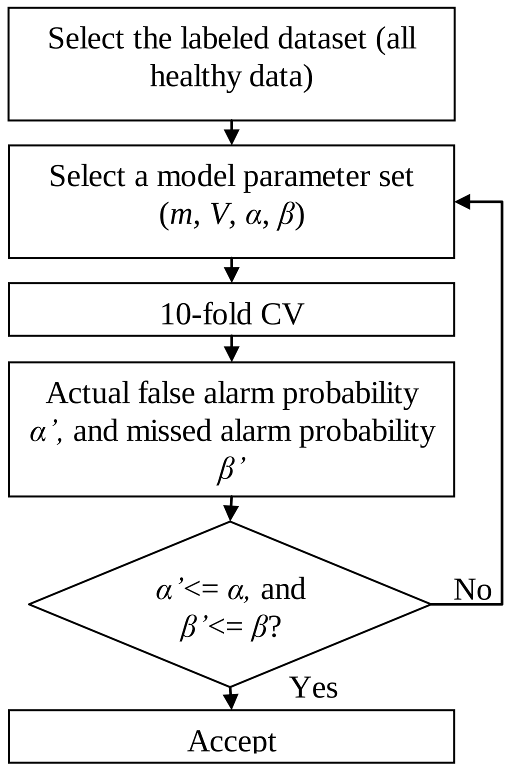

Figure 5: Model parameter selection procedure by 10-fold CV.

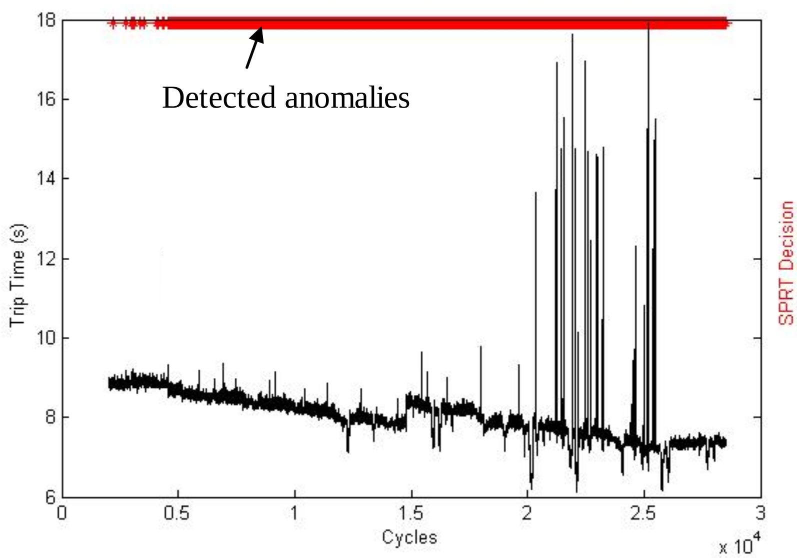

Figure 6: Anomalies detected by SPRT.

RUL Prediction for PPTC Resettable Fuses

The RUL of a product can be predicted by several categories of methods, including data-driven methods, model-based methods, and the fusion of data-driven and model-based methods. In this study, the data-driven method was used because it is simple and can capture the complex features of a product without knowledge of the failure mechanisms. Some data-driven RUL prediction models directly estimate the failure of a product by forecasting the precursor parameter until the failure criteria are reached. The precursor parameter can be considered to be a time-series signal, and then the forecasting methods can be used to estimate the future value of the parameter.

We used the ARIMA model to predict RUL for PPTC resettable fuses because the linear trending is exhibited in the data (Figure 2). The ARIMA model is one of the most widely used linear models in time series forecasting, and was introduced by Box and Jenkins [10]. In this model, a time series *y* can be modeled as linear function of current and past values. The *ARIMA(p,d,q)* model can be identified as the *AR(p)* model, the *MA (q)* model, or the *ARMA(p,q)* model, in which *p* and *q* are the order of the process, meaning that the future value is a combination of *p* and *q* previous values from now, and *d* is the time of data deference, which is for a non-stationary process.

If *y* is a stationary process, for an autoregressive process of order *p*, $y \sim AR(p)$ or $ARIMA(p,0,0)$:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + \varepsilon_t \tag{10}$$

where a_i are coefficients, $\varepsilon \sim N(0,\sigma^2)$;

For a moving average process of order *q*, $y \sim MA(q)$ or $ARIMA(0,0,q)$:

$$y_t = \varepsilon_t + b_1 \varepsilon_{t-1} + b_2 \varepsilon_{t-2} + \ldots + b_q \varepsilon_{t-q} \tag{11}$$

where, b_i are coefficients, $\varepsilon \sim N(0,\sigma^2)$;

For the combination of the *AR* and *MA* models, or the *ARMA* process, $ARMA(p,q)$ or $ARIMA(p,0,q)$:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + \varepsilon_t + b_1 \varepsilon_{t-1} + b_2 \varepsilon_{t-2} + \ldots + b_q \varepsilon_{t-q} \tag{12}$$

If *y* is a non-stationary process, conduct data difference for $d >= 1$ times, which is denoted as y^d_t, meaning conducting $(y_t - y_{t-1})$ for *d* times, then finding (p,q).

The whole ARIMA(p,d,q) modeling procedure includes model identification, estimation, and model validation. Model identification defines the values of *p, d,* and *q* by several plots, which are plots of data, autocorrelation function (ACF) plots, and partial autocorrelation function (PACF) plots. Figure 7 illustrates the procedure of model identification.

In model estimation the identified ARIMA model is fitted to the training data and estimates the parameters (coefficients). The standard errors from the estimates and 95% confidence limits for the parameters are also calculated to determine whether the model is appropriate. In model validation, the residuals between the estimation and actual data should be white noise (or independent when their distributions are normal) drawings from a fixed distribution with a constant mean and variance. Otherwise, a better model should be considered.

The prediction based on the data shown in Figure 2 is used as a case study. The data is first re-sampled by averaging the trip time every 100 cycles, as shown in Figure 8, from which we can see that feature of the data is maintained. Based on the anomaly detection results, we conducted the RUL prediction from the first detected anomalies. Here, two RUL predictions are demonstrated as general examples. One was predicted from the 101st data point to the 150th data point based on the previous 100 data points. In this case, the ARIMA model was identified as ARIMA (0,2,2), and the prediction was shown in Figure 9. From the prediction, the earliest possible failure will be at the 118th point (11800 cycles, 95% CI lower limit), and the actual first intermittent failure is at the 124th point (12400 cycles). In the second case, trip time was predicted from the 191st data point to the 240th data point based on the previous 190 data points. The model was identified as ARIMA (0,2,1), and the prediction was shown in Figure 10. The ARIMA model is updated and new coefficients were calculated. Based on the prediction, the possible earliest failure would be at the 191st point (19100 cycles, 95% CI lower limit). The prediction line showed that the failure would be at the 198th point (19800 cycles), which is earlier than the actual failure at the 205th point (20500 cycles).

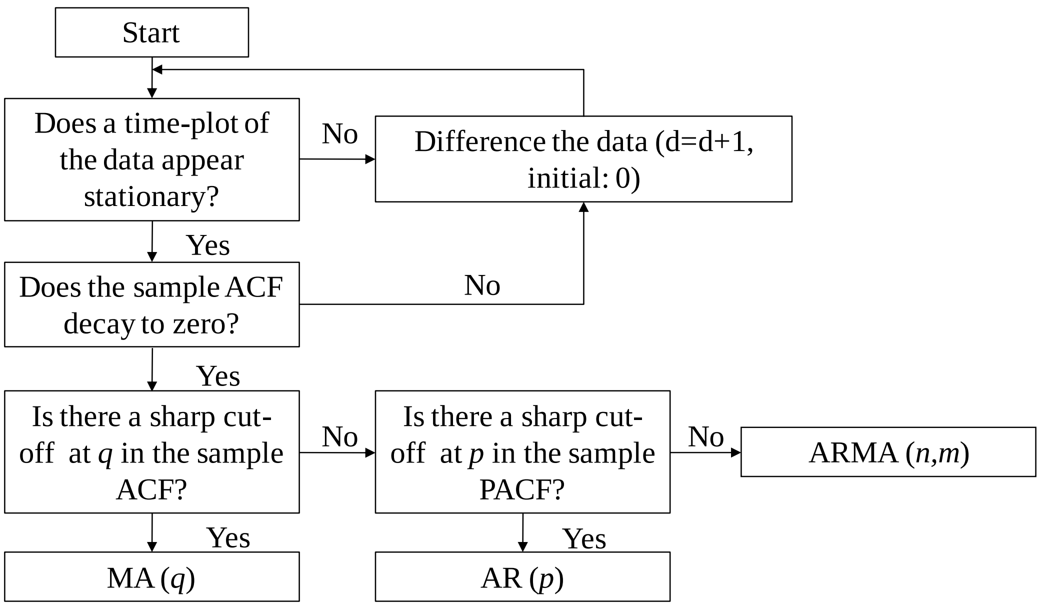

Figure 7: ARIMA model identification procedure.

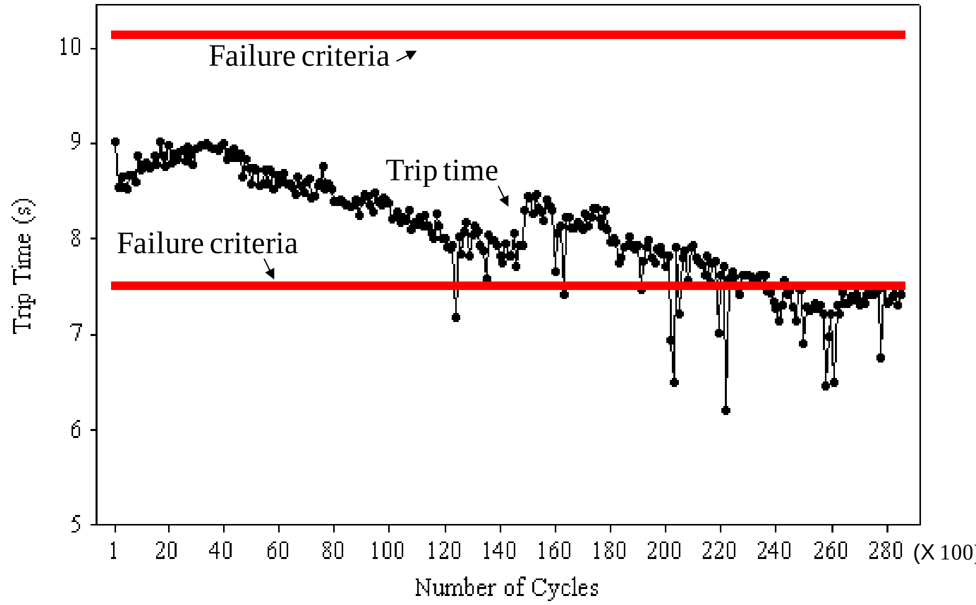

Figure 8 Trip time data re-sample.

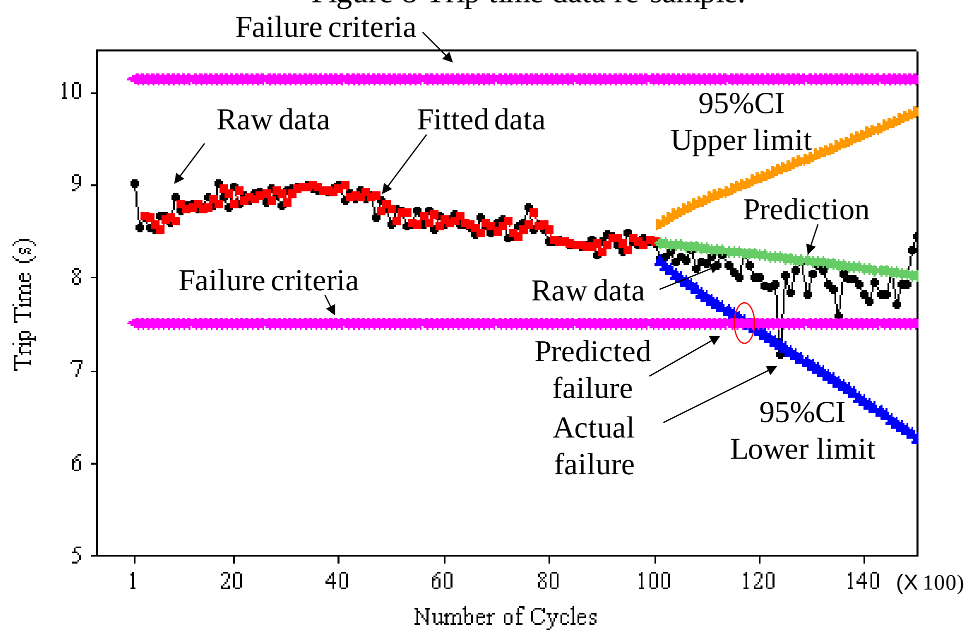

Figure 9: Prediction from the 101st to 150th points based on the previous 100 points

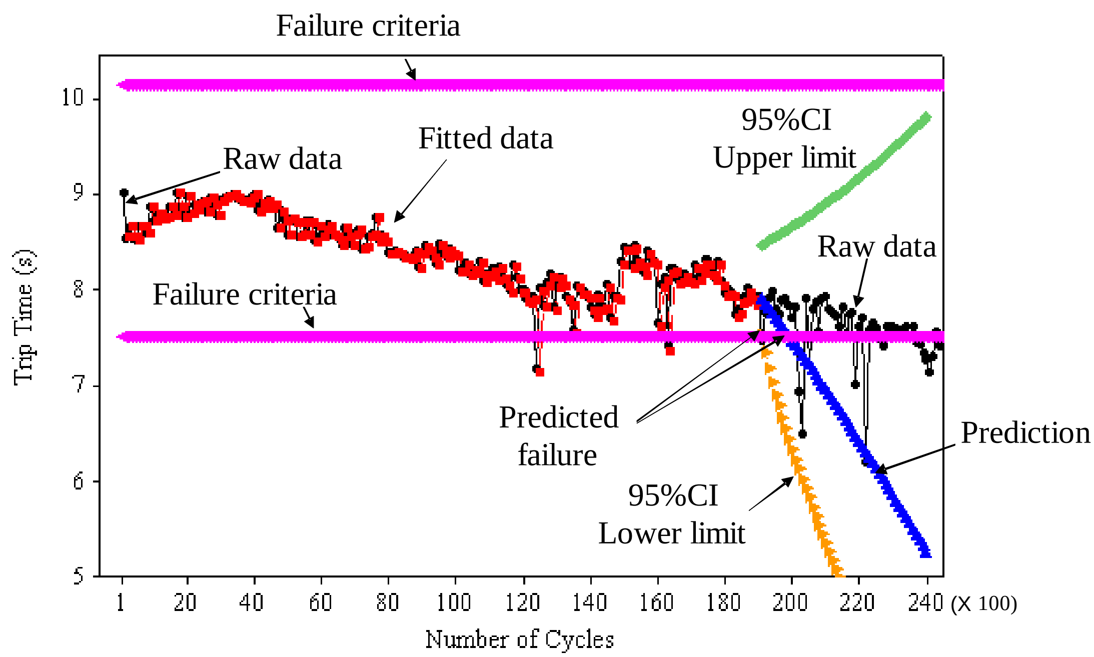

Figure 10: Prediction from the 191st to 240th points based on the previous 190 points.

Discussion and Conclusion

The RUL of PPTC resettable fuses can be predicted in situ by estimating the future value of the monitored failure precursor parameters, such as the trip time with cycles. Anomaly detection can be conducted in situ by integrating a cross-validation technique into traditional SPRT to provide in-time or advance alarms. The integration of the CV technique into SPRT (called CV-SPRT) solves the difficulty in the model parameter selection for SPRT so that proper model parameters for SPRT can always be found without the requirement of experience. The CV-SPRT is not limited to anomaly detection of PPTC resettable fuses; it can be extended as a standard method for any in-situ anomaly detection when the distribution of the data is available or predictable.

A failure precursor parameter can be treated as a time-series signal, which can be predicted by a mathematical model. The selection of the prediction model should consider the features of the signals and the properties of the model. The autoregressive integrated moving average (ARIMA) model is one linear model. It was used in this paper to predict the future values of the failure precursor parameters due to their exhibiting linear trends in the trip time signals. RUL can then be estimated when the predicted future values cross the failure criteria. Based on the demonstrations of RUL prediction of PPTC resettable fuses, ARIMA is a practicable method for failure prediction for the linear failure precursors. This is because the ARIMA model assumes that the time-series signals are linear. Other techniques, such as neural networks, could be incorporated with ARIMA to capture both the linear and nonlinear components of the signals included in the failure precursor parameters.

References

[1] Tyco Electronics, “Fundamentals of Polyswitch Overcurrent and Overtemperature Devices, Technical paper,” 2008.

[2] S. Cheng, M. Azarian, and M. Pecht, “Sensor System for Prognostics and Health Monitoring,” Sensors, No. 10, pp.5774-5797, 2010.

[3] M. Pecht, *Prognostics and Health Management of Electronics*. New York: Wiley-Interscience, 2008.

[4] S. Cheng, K. Tom, and M. Pecht, “Failure Precursors for Polymer Resettable Fuses,” IEEE Trans. on Devices and Materials Reliability, Vol.10, Issue.3, pp.374-380, 2010.

[5] UL 1434. Standard for Safety for Thermistor-Type Devices, 2002.

[6] S. Cheng, K. Tom, L. Thomas, and M. Pecht, “A Wireless Sensor System for Prognostics”, IEEE Sensors Journal, Volume 10, Issue 4, pp. 856 – 862, 2010.

[7] A. Wald, *Sequential Analysis*, John Wiley and Sons, New York, NY, 1947.

[8] A. Wald, Sequential Test of Statistical Hypotheses, The Annals of Mathematical Statistics, Vol. 16, No. 2, pp. 117-186, 1945.

[9] H. Sohn, D. Allen, K. Worden, and C. Farrar, “Statistical Damage Classification Using Sequential Probability Ratio Tests,” Structural Health Monitoring, Vol. 2(1), pp. 57–74, 2003.

[10] G.E. Box, G.M. Jenkins, and G.C. Reinsel, *Time Series Analysis Forecasting and Control, 4th edition*, John Wiley and Sons, New York, 2008.

[11] K.C. Gross and W. Lu, “Early Detection of Signal and Process Anomalies in Enterprise Computing Systems,” Proc. 2002 IEEE Int’l Conf. on Machine Learning and Applications (ICMLA), Las Vegas, NV, 2002.

[12] J.P. Herzog, S.W. Wegerich, R.M. Singer, and K.C. Gross, “Theoretical Basis of the Multivariate State Estimation Technique (MSET),” Argonne National Laboratory NT Technical Memorandum No. 49, 1997.

[13] A. Miron, “A Wavelet Approach for Development and Application of A Stochastic Parameter Simulation System,” PhD Dissertation in Nuclear Engineering of University of Cincinnati, 2001.