Aircraft APU Starter Health Monitoring and Failure Prognostics

by

Qingfeng LOU

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Applied Science

in

Aerospace Engineering

Carleton University
Ottawa, Ontario

Abstract

Auxiliary power unit (APU) is a gas turbine engine on aircraft that provides energy for functions other than propulsion. Its starter is a crucial component that outputs assistant power to support the APU starting process. Starter performance degradation significantly impairs the whole APU life and raises risks for the aircraft flight. However, the current maintenance policy for the starter is still “run it till it breaks”. An effective technique for the starter diagnostics and prognostics has not been reported yet. The aim of this thesis is to propose a framework for enabling the online detection and prediction of starter degradation. For this purpose, the thesis makes use of a dataset containing information about 52 APU “inability to start” failure events that were collected from actual aircraft operations over a period of ten years. Through the establishment of the relationship between the starter degradation and gas turbine engine starting performance, 13 of these 52 failures were identified as being caused by the starter degradation. Once this determination has been made, an online classifier based on moving autocorrelation is designed to detect the initial phase of degradation for each failure. Finally, a particle filtering based approach with an associated system state model is proposed to achieve the fault diagnostics and failure prognostics. The results demonstrate that a condition based maintenance program for the APU starter can be implemented to avoid unnecessary economic losses and to enhance aircraft operating safety.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor Dr. LIU Jie for his enthusiastic supervision and patient guidance. He is the best supervisor I had ever met, with encourage, enthusiasm, and immense knowledge. He supported and guided me throughout the entire master program. Without him, it would not have been possible to have this thesis. Also, I would like to sincerely thank our research collaborator Dr. YANG Chunsheng from National Research Council Canada for his insights and nice helps in this two-year project.

I would also like to thank LI Geng and YANG Jiawei, my best friends and excellent aero engine engineers, for their tremendous technical suggestions in the airline maintenance field. Many thanks should also be given to Instructor YU Hai from the CAE flight training center, Instructor HOU Gang from the Airbus flight training center, and Captain YU Zhou from China southern airline, for their tremendous technical assistance in airline operation field. In addition, I wish to thank Dr. Katie Macnamara for her nice suggestions in thesis preparation.

In the meantime, my loves and thanks should be going to my dear parents and my sweet wife ZHAN Zhao for providing me endless housework that prevents me from being bored, also for her tremendous technical consultations in gas turbine engine theory.

Contents

Figures

Tables

Nomenclature

C_k:	Particle filtering state to describe discrete fluctuations of $\overline{EP}_k$
c_p:	Specific heat at constant pressure
c_{pa}:	Specific heat at constant pressure of ambient air
c_{pg}:	Specific heat at constant pressure of combustion gas
cyc:	Cumulative count of APU starting cycles
EGT:	Exhaust gas temperature (or turbine outlet temperature)
EGT_c:	Corrected EGT_{raw} based on ambient conditions
EGT_{peak}:	Peak value of EGT during starting process
EGT_{raw}:	Raw EGT data measured from aircraft operation
EGT_{stable}:	EGT when air conditioning is enabled after starting with $100\%\ N$
EP:	Corrected EGT_{peak}
$\overline{EP}_k$:	Particle filtering state to describe starter health condition
ES:	Corrected EGT_{stable}
f:	Fuel/air ratio
h_{op}:	Cumulative count of APU operating hours
ISA:	International standard atmosphere, at sea level $T_{ISA} = 288.15\ K$, $P_{ISA} = 101.325\ kPa$
m:	Air mass flow
N:	Rotational speed

N_c:	Corrected N_{raw} based on ambient conditions
N_{peak}:	N corresponding to EGT_{peak}
N_{raw}:	Raw N data measured from aircraft operation
NP:	Corrected N_{peak}
$P_{starter}$:	Starter assistance power
$P_{w_{net}}$:	Engine net power output
$P_{w_{net}}$:	Engine external power load
P_1:	Ambient air pressure (or compressor inlet pressure)
P_2:	Compressor outlet pressure
P_3:	Turbine inlet pressure
P_4:	Turbine outlet pressure
R:	Gas constant
R_X:	Theoretical autocorrelation
R'_X:	Theoretical autocovariance
RUC:	Remaining useful cycles
Sn:	APU serial number
S_X:	Power spectral density function
t_{start}:	Time duration of starting process
T_1:	Ambient air temperature (or compressor inlet temperature)
T_2:	Compressor outlet temperature
T_3:	Turbine inlet temperature
T_4:	Turbine outlet temperature

ts:	Corrected t_{start}
V_X:	Experimental autocorrelation
w_k^i:	Importance weight
x_k^{i-}:	Original particles of particle filtering at time step k
x_k^{i+}:	Resampling particles at time step k
y_k:	Measurement at time step k
Y_k:	All available measurements at time step k.
γ:	The ratio of specific heat
ΔT_{12}:	Compressor temperature rise
ΔT_{34}:	Turbine temperature drop
η_c:	Engine compressor efficiency
η_m:	Mechanical efficiency of the compressor-turbine combination
Θ:	T_1/T_{ISA}
λ_k:	Particle filtering state to describe the exponential growth rate
μ_{nor}:	Mean value of condition monitoring data in the normal phase
σ_{nor}:	Standard deviation of condition monitoring data in the normal phase
τ:	Time difference variable of R_X

Chapter 1 Introduction

1.1 Overview

1.1.1 APU starter

Auxiliary power unit (APU) is standard equipment on today's civil aircraft. It is a small gas turbine engine that makes the aircraft independent of ground power supplies. In normal operation, APU works during the phases when the aircraft's main engines are shut down (e.g. cockpit preparation), to support the electrical and air conditioning systems. The other primary function of APU is to provide pneumatic power for the main engines starting. In adverse condition operation (e.g. high altitude/high weight take-off) or in abnormal operation (e.g. engine out in flight), APU enhances the electrical margin and relieves the working engines' load [1].

APU starter is one of the most crucial components of APU. During the starting process, the starter accelerates APU to a high rotational speed to provide sufficient air compression for self-sustaining operation. When the starter performance gradually degrades and its output power decreases, either the APU combustion temperature or the surge risk will increase significantly. These consequences will then greatly shorten the whole APU life and even result in an immediate thermal damage. Thus the APU

starter degradation can result in unnecessary economic losses and impair the safety of airline operation. The current maintenance strategy in civil aviation for the APU starter, however, is still the breakdown maintenance. This strategy means that losses caused and risks incurred by the starter degradation cannot be avoided.

1.1.2 Condition-based maintenance

Maintenance is essential in keeping machines and equipment safe and reliable. The most early and basic maintenance strategy is *breakdown maintenance*. This strategy basically lets the target equipment 'run it till it breaks'. This means that no action will be taken until the equipment fails. The later developed maintenance strategy is *preventive maintenance*. In this strategy, actions are performed at periodic intervals or according to machine-run based schedules. The preventive maintenance can effectively prevent the degradation of the equipment. However, one major drawback to this maintenance strategy is that it ignores the actual health status of the equipment. Thus catastrophic failures are still likely to occur and some costly as well as unnecessary maintenance actions are performed.

Condition-based maintenance (CBM) is a more recently developed strategy that recommends maintenance decisions based on the information collected through system condition monitoring and equipment failure prognostics. A CBM, if well established and properly implemented, can significantly reduce the expenditures by avoiding unnecessary maintenance operations while eliminate catastrophic equipment

failures and maintain the performance in a satisfactory level [2].

A CBM framework usually consists of three key steps: (1) Data acquisition, (2) Data processing, (3) Decision-making. Data acquisition is the collection of useful data from targeted equipment for the purpose of CBM. Data processing involves the analysis and interpretation of data for the purpose of extracting features related to equipment health conditions. Decision-making consists of two main aspects, diagnostics and prognostics. Diagnostics monitors the historical data and current status of the targeted equipment to detect the faults occurrence. Prognostics anticipates future status changes and prevents faults before they occur.

The research in this thesis focuses on the implementation of an entire CBM framework on the APU starter to support the corresponding practical civil aviation maintenance programs.

1.2 Literature review

1.2.1 Gas turbine engine theory

The basic theories of gas turbine engines formed by thermodynamics and aerodynamics have been well developed for decades. Since APU is a typical type of gas turbine machine delivering shaft power, the gas turbine engine theories are well applicable to it. The text books written by Saravanamuttoo [3] and Kerrebrock [4]

have offered good introductions to Gas turbine engines, covering theories from main components design to overall operation principles. More precisely, Walsh [5], focused on gas turbine engine performance, presenting general performance calculation methods for evaluation of engines in balanced and transient states as well as engines specifically engaged in the starting process. The book published by Rolls-Royce [6], by contrast, focused on gas turbine engine structures, offering detailed mechanism information of engine starting system.

While the texts mentioned above primarily discuss engines operating in normal states, the abnormal states of different components have been investigated in a number of research works as well. Aker in [7] and Saravanamuttoo in [8] for instance, have evaluated the engine performance degradation caused by compressor fouling using computer simulation techniques; Hanachi, *et al*. in [9] proposed a procedure for diagnosing compressor blade cracks based on detrended fluctuation analyses (DFA) of an experimental dataset; And Kim in [10] applied a fuzzy clustering method to engine operating data in order to diagnose faults in fuel supply systems. However, little attention has thus far been paid to gas turbine engine starter degradation, and the impact of starter degradation on engine performance has rarely been discussed in literatures.

As for a particular type of APU installed in civil aircraft, generally, the user level information (for airline) about the structure and operation procedures are provided in the flight crew operating manuals (FCOM) [11] and the aircraft maintenance manuals (AMM) [12]. However, the design level information (from manufacturer) such as the

performance charts and engine simulation models are normally unavailable to the public. In this thesis, a set of data collected during 10 years of actual airline operations has been used to evaluate the particular APU's performance.

1.2.2 Diagnostics

Machinery diagnostics is a posterior data analysis that uses historical and current system state to achieve early fault detection and localization. The conventional diagnostics approaches rely on human expertise by utilizing feature extracting tools such as power spectrum [13], cepstrum [14], wavelet phase [15], etc. However, the manual diagnostics approaches are usually time consuming and unreliable when multiple features are involved or when the data are interfered by noises. Alternatively, the progressive diagnostics approach can achieve automatic fault detection. Based on different emphases, the automatic diagnostics approaches can be generally classified into three categories: statistical approaches, artificial intelligence (AI) approaches, and physical-based approaches [16].

Cluster analysis is a widely used statistical approach that classifies signals into different fault and health groups on the basis of the similar features possessed by the signals. One way of cluster analysis is to represent the similarity among different signal groups with the distance measures derived from certain discriminant functions in statistical pattern recognition [17]. The commonly used distance measures includes Euclidean distance [18], Kullback-Leibler distance [19], Bayesian distance [20] etc.

Another way of cluster analysis is to determine the boundary curves that separate different signal groups. Support vector machine (SVM) is a powerful technique applied for such a boundary determination. The basic principle is to separate an n-demensial data space by a $n-1$-demensianl hyperplane that creates the maximum margin between two adjacent groups. Other than cluster analysis, hidden Markov model (HMM) can also be used for fault classification [21]. It assumes the time-frequency features extracted from the original signals to be Markov processes with hidden states. The targeted component's health condition are then represented and classified by evaluating these hidden states [22]. Generally, all the statistical approaches have a main limitation that they cannot be used for time-varying systems, since they rely on the statistic measurements.

Neural network (NN) and fuzzy logic are two useful AI approaches. The NN is a computational model that mimics human brain nervous system that is capable of machine learning and pattern recognition. It has a complex layer structure consisted of simple processing nodes that can model complex non-linear function with multiple inputs and outputs [23]. Fuzzy logic deals with system uncertainties and ambiguities in a way that mimics human reasoning [24]. It starts from highly formalized insights about the structure of categories, and then formulates expert knowledge in a linguistic form [25]. One example of the AI approach applications is given in [26], in which a fault diagnostic scheme is proposed to integrate NN and fuzzy logic for an assessment of bearing health condition. Generally, the AI approaches have two main limitations. Firstly, it is difficult to have physical explanations on the training models. Secondly, a

comprehensive amount of data is required for the training process which is often unavailable for certain cases.

Physical-based approaches incorporate physics understanding and explicit mathematical model of targeted system into fault diagnostics. With the established model, the estimation tools such as Kalman filter, parameter estimation can be applied to implement the diagnostics. One example of the applications is given in [27], in which a model-based diagnostics system has been built to represent the rolling element bearing faults. Another example is given in [28] in which the fault diagnostics of cracked rotor shafts has been implemented. Generally, with the accurate and reliable model, the physical model based approaches can be more effective than other model-free approaches. However, one main limitation is that each particular targeted system requires its specific model. For some complex systems, such models are very difficult or even impossible to be established.

1.2.3 Prognostics

Machinery prognostics uses current and previous system states to predict a dynamic system's future states. Reliable forecast information can be used to schedule repairs and maintenance in advance and provide an alarm before faults reach critical levels so as to prevent system performance degradation, malfunction, or even catastrophic failures [29].

In general, prognostics can be conducted using either data-driven methods or

physics-based approaches. Data-driven prognostic methods use pattern recognition and machine learning techniques to detect changes in system states [30], [31]. The classical data-driven methods for nonlinear system prediction include the use of stochastic models such as the autoregressive (AR) model [32], the threshold AR model [33], the bilinear model [34], the projection pursuit [35], the multivariate adaptive regression splines [36], and the Volterra series expansion [37]. In the last decade, interest in data-driven system state forecasting that focuses on the use of flexible models such as various types of neural networks (NNs) [38]，[39] and neural fuzzy (NF) systems [40], [41] has grown. Data-driven prognostics methods rely on past patterns of the degradation in similar systems to project future system states. Their forecasting accuracy depends on not only on the quantity but also on the quality of historical system data, which could be a challenging task in many real applications [42]，[43]. Another principal disadvantage of data-driven methods is that the prognostics reasoning process is usually opaque to users [44]. As a result, data-driven methods are often not suitable for some applications where forecast reasoning transparency is required.

Physics-based approaches typically involve building models (or mathematical functions) to describe the physics of the system states and failure modes. These approaches incorporate a physical understanding of the system into the estimation of system state and/or remaining useful life (RUL) [45]，[46]，[47]. Physics-based approaches, however, may not be suitable for some applications where the physical parameters and fault modes may vary under different operation conditions [48]. On

the one hand, it is usually difficult to tune the derived models to accommodate time-varying system dynamics. On the other hand, physics-based approaches cannot be used for complex systems whose internal state variables are inaccessible (or hard) to direct measurement using general sensors.

The Bayesian approaches are particularly well suited to solve dynamic state-estimation problems since they incorporate process data into an optimum state's estimation by considering the likelihood of observations [49]. Particularly, the particle filtering (PF), also known as sequential Monte Carlo (SMC), provides a solid and consistent theoretical framework for handling nonlinear/non-Gaussian system state models. The PF based approaches, however have been used only in a few prognostics applications, such as the prediction of battery RUL [50] or the monitoring of crack growth in engine blades [51]. One of the common limitations for PF prediction is that no future measurements are available to correct for model inaccuracies. In that sense, an accurate system state model of PF based on a physical understanding of the target equipment is necessary for reliable and robust predictions.

1.3 Objectives

The related researches of this thesis are Letourneau 2008 [52] and Yang 2010 [53] on behalf of the National Research Council Canada (NRC). Letourneau *et al*. in [52] presented a classification system based on clustering and support vector machine (SVM) for the starter failure estimations. Yang *et al*. in [53] presented the APU failure

mode and effect analysis. Through an investigation of maintenance records, the paper validated all the possible component faults that can lead to the “APU inability to start” failure effect.

This thesis aims to develop a framework that supports a complete CBM framework for the APU starter in practical civil aviation field. The proposed framework will cover the methodologies for implementing all the three CBM key steps, from data acquisition to data processing to decision-making. This thesis is new in the following aspects:

1. It first time establish the physical relationship between engine performance and the starter health condition based on gas turbine engine theory. With these relations established, different APU failure modes can be identified based on data patterns of collected engine parameters. Also the starter degradation can be indicated by these parameters.
2. It designs an online classifier based on autocorrelation by considering engine parameters as random processes. With this classifier thus designed, the initial phase of starter degradation can be effectively detected.
3. It develops a PF approach with the associated system state model of the APU starter. This PF approach can remove the inherent uncertainties in the data measurements, automatically update the system state model parameters, and dynamically achieve diagnostics and prognostics at each starting cycle for supporting the CBM decision-making.

1.4 Thesis Outline

The reminder of this thesis is organized as follows.

Chapter 2 focuses on the data acquisition. First, relevant background information about the APU and starting procedure is introduced. Then the dataset employed in the thesis is interpreted. Base on the interpretation, several essential engine operating parameters are collected for subsequent technology development.

Chapter 3 presents the data processing. The engine parameters are corrected with respect to different ambient conditions. After that, different data patterns are identified from these corrected parameters.

Chapter 4 discusses gas turbine engine theory. The physical relationship between engine performance and the starter health condition are established. With this physical relationship, this chapter establishes a method by which the starter health condition can be indicated by the collected engine parameters.

Chapter 5 presents the diagnostics of decision making. The statistical properties of the dataset are examined with the methods of kernel smoother and moving average. After that, an online classifier based on autocorrelation is designed to detect the starter degradation.

Chapter 6 implements the diagnostics and prognostics of decision making. A PF estimator is developed with the proposed system state model. This PF estimator can provide optimum estimation of starter current condition and the future degradation trend based on available engine parameter measurements.

Finally, Chapter 7 presents the conclusions of this thesis and some future research topics.

Chapter 2 Data acquisition

Data acquisition is the process of collecting useful data from target equipment. This process is an essential step in implementing a condition-based maintenance (CBM) program for machinery fault diagnostics and failure prognostics.

In this chapter, Section 2.1 introduces the necessary background information about APU operation and staring process. Section 2.2 introduces the data acquisition with necessary interpretations.

2.1 APU structure

The target equipment in this thesis is the Honeywell 331 series APU installed on Airbus A310 [11]. It is a self-contained unit installed in the fireproof compartment located in the aircraft fuselage tail cone. As shown in the Figure 2.1 (a), APU is a small gas turbine engine that delivers mechanical shaft power for the aircraft electrical network and produces bleed air for the main engines starting and air conditioning systems. The APU contains the following main components:

- Power section: A two-stage centrifugal compressor driven by a three-stage axial turbine through the main shaft of a single-spool configuration.
- Load compressor: A single-stage centrifugal compressor directly driven by the power section. The bleed air is delivered from this component to the aircraft pneumatic system.

- Accessory gearbox: A shaft power delivery mechanism driven by the power section directly, carrying the Fuel Control Unit, lubrication pumps and AC generator.
- Starter: A series-wound DC electric motor that automatically varies its speed according to the load, with speed increasing as load decreases and vice versa. During starting, the starter receives power from the aircraft battery, and then drives the APU power section through gearbox.
- Electronic Control Box (ECB): A full-authority digital electronic controller that performs the bulk of the APU system logic for all operations. These operations include starting and shutdown sequence, as well as the monitoring of rotational speed, temperature, bleed and electrical load.
- Air intake flap: It is an electrically operated flap that allows external air to reach the compressor inlet.

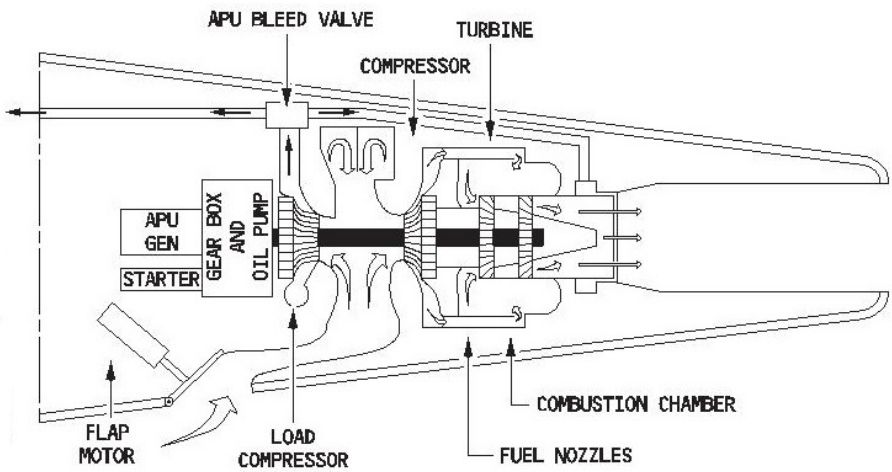


Figure 2.1 APU schematic [11].

Figure 2.2 is the flight phases diagram described by the Airbus FCOM. During the flight operation, the APU is normally started on the ground in the phase of cockpit preparation before main engine starting or after touchdown (Phases 1, 9 and 10). The APU is usually shut down in the other phases when the main engines are working.

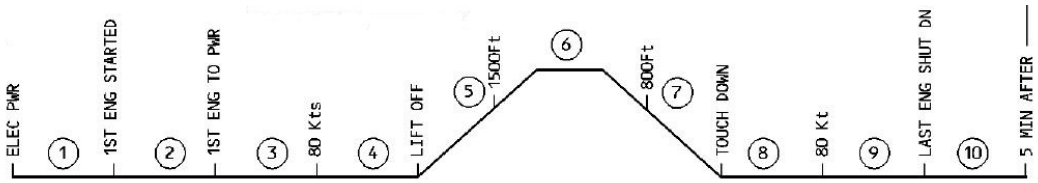


Figure 2.2 Airbus flight phases diagram [11].

APU starting is an automatic sequence performed and monitored by the ECB. Once the pilot in the cockpit selects the master switch ON and presses the APU START pushbutton switch, the starting process is initiated. During starting, the ECB controls fuel flow by considering the rotational speed (N), Exhaust Gas Temperature (EGT) and air inlet conditions. This logic flow is schematically shown in Figure 2.3 and the key phases are listed as follows:

- Dry cranking: After the air intake flap reaches its full open position, the starter is energized to provide initial rotational speed to the APU spool with no fuel supply.
- Light-off: At 7% N, fuel is metered to the combustor, and the igniters are energized to ignite locally within the combustor.
- Acceleration: the APU spool is continually accelerated by the thermal power

of fuel combustion and the shaft power of the starter. N increases steadily along with the fuel supply.

- Starter cutoff: At 50% N, the starter is declutched and de-energized. After that, the APU is self-sustained and accelerated by thermal power only.
- Starting completed: At 95% N, the APU is available to provide electrical load and bleed load to the aircraft. When the starting process is completed, ECB ensures the APU to work stably at 100% N.

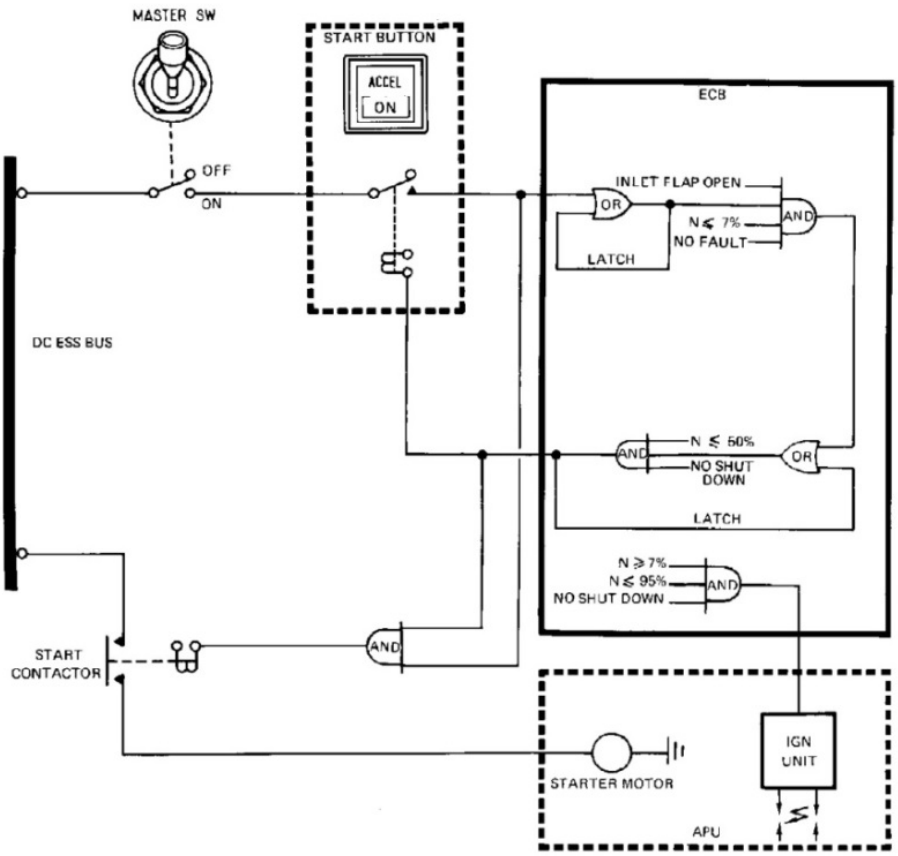


Figure 2.3 APU starting logic schematic [11].

Figure 2.4 records the key APU parameter profiles of starting process from A320

cockpit in an actual flight. The A320 APU has a structure and starting logic that is quite similar to that of A310 APU. Although the numerical values have certain differences, the general trends of these parameter profiles are similar. As can be seen in the figure, during the starting process, N increases evenly from 0% to 100%. EGT increases rapidly after light-off, reaches the peak before the starter's declutching and decreases gradually to a stable value as N reaches 100%. In the subplot (d), the point where EGT reaches its peak value is denoted as (N_{peak}, EGT_{peak}). This point is a distinctive feature that describes the whole profile of EGT against N.

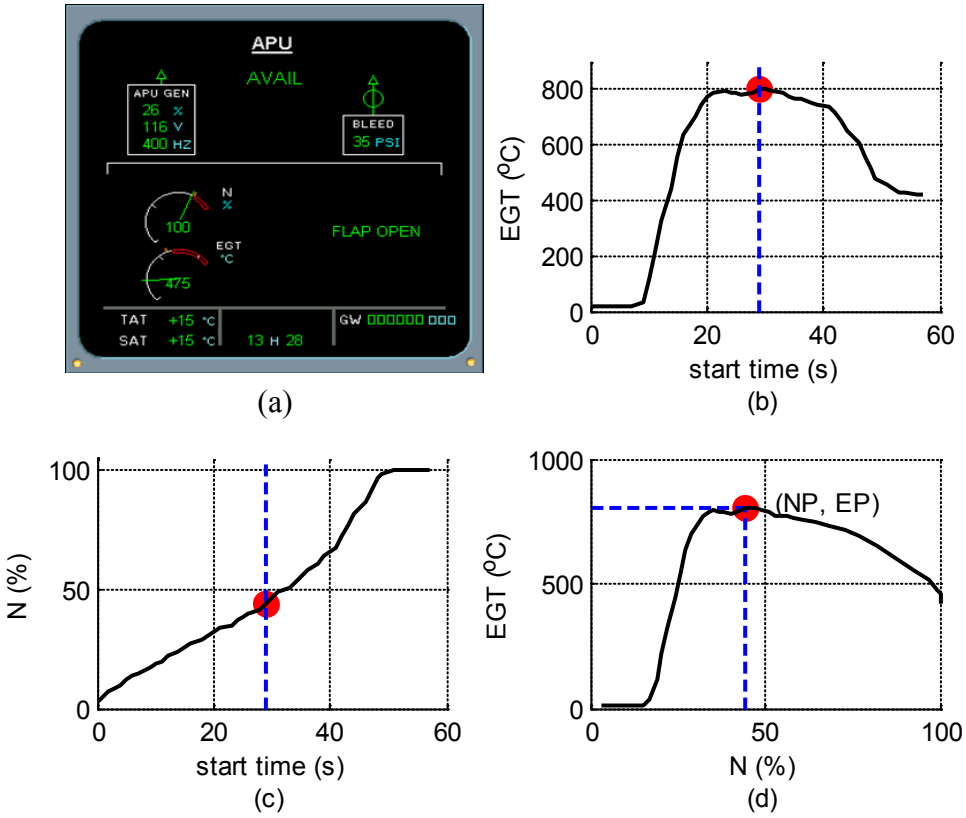


Figure 2.4 APU starting parameters (a) cockpit display, (b) N against $start\ time$, (c) EGT against $start\ time$, (d) EGT against N.

2.2 Dataset acquisition

Data collected for a CBM framework can be categorized into two main types: the *condition monitoring data* and *event data*. Condition monitoring data are the measurements related to the health condition of target equipment, while event data include the information about what has happened to target equipment (e.g., maintenance, breakdown, etc.). In this thesis, both of the two data types are involved in data acquisition.

2.2.1 Condition monitoring data

The dataset in this thesis was derived from the flight records of the Air Canada fleet of 35 Airbus-A310 aircrafts. These records contain APU operating data for 161,000 commercial flights over more than a 10-year period.

Table 2.1 demonstrates part of the dataset for an easy overview. Each row of the table contains data from one independent APU starting cycle. Based on the information of APU operations introduced previously, 10 variables from 70 have been picked out as the *condition monitoring data* because they are related to the APU starting conditions, while others such as oil temperature and bleed valve angle are either too vague to be interpreted, or are collected after the starting process. These 10 variables are listed below alongside their respective annotations.

- $Info$ Flight information (flight number, date and route, etc.).
- Sn APU serial number.
- h_{op} Cumulative count of the APU operating hours.
- cyc Cumulative count of the APU starting cycles.
- T_1 Ambient air temperature.
- P_1 Ambient air pressure.
- EGT_{peak} Peak value of exhaust gas temperature in starting process.
- N_{peak} Rotational speed at the moment of EGT_{peak} occurrence.
- t_{start} Time duration of starting process.
- EGT_{stable} Exhaust gas temperature when air conditioning is enabled after starting with 100% N.

2.2.2 Event data

A comprehensive search of the maintenance reports in the same period of the dataset reveals that 83 component replacements have occurred due to the failure effect: "APU inability to start". However, neither the exact failure modes nor the name of faulty component were made available in these reports. That is, APU "inability to start" failures were recorded, and a faulty component was replaced, but neither the failure modes associated with these failures were nor the components involved were specifically identified. Thus the failure modes could have included hot-starting, hung-starting, ignition failed, or something else. And the faulty components could

have included the starter, the fuel management controller, the igniters, or something else.

The problem of identifying the exact failure mode and faulty component will be handled later in Chapter 4 after the gas turbine theory of starting process is discussed. But at this moment, the 83 component replacement events are treated equally as *event data*, and are named as *reset events* in the following context for simplicity. This terminology is used because it is after a component replacement event that both the h_{op} and cyc in the dataset stop their cumulative counts and restarted from zero, as indicated in Table 2.1. This means that the condition monitoring data always reflect a reset event after a component has been replaced

Furthermore, only the condition monitoring data from the last 250 APU operating hours prior to the reset events were made available. This is because it is during this 250-hour period that all the relevant degradation information about the affected APU component is believed to be recorded. Another terminology, *group*, is used in this thesis that all of the condition monitoring data corresponding to one reset event are regarded as one *group*. That is to say, each group corresponds to one APU's record that a certain component has experienced degradation which finally led to an APU "inability to start" failure effect. After data with format and sensor errors have been eliminated and groups with insufficient length have been removed, 52 valid groups are ultimately generated from the entire dataset.

Table 2.1 Overview of the dataset.

$Flight\ info$	$Dest \leftarrow Dept$	EGT_{peak}	$\cdots$	Sn	h_{op}	cyc	RUC	
20SEP2002:11:54:26	YYC-YYC	527	$\cdots$	795	1986	3150	394	
20SEP2002:17:07:26	YVR-YYC	538	$\cdots$	795	1988	3151	393	
20SEP2002:19:53:44	YYC-YVR	532	$\cdots$	795	1988	3152	392	Group 1
20SEP2002:22:35:48	YVR-YYC	531	$\cdots$	795	1989	3153	391	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
19DEC2002:19:27:13	YYZ-PBI	872	$\cdots$	795	2235	3542	2	
20DEC2002:00:58:44	YEG-YYZ	792	$\cdots$	795	2236	3544	0	Reset event
25JAN2005:16:10:11	YYZ-YUL	538	$\cdots$	876	4	4	183	
25JAN2005:20:07:20	YYZ-YOW	533	$\cdots$	876	5	7	180	
26JAN2005:21:58:37	YVR-YYZ	498	$\cdots$	876	18	15	172	Group 2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		

As mentioned in Section 2.1, during one starting cycle, the APU starter only works for a short period from $0\%\ N$ to $50\%\ N$ regardless of how long the APU as a whole works after starting. Therefore, it is reasonable to use starting cycles as time units to measure the starter's life span rather than to use the operating hours.

In order to effectively combine the event data and condition monitoring data together to describe the underlying component degradation, the term, *remaining useful cycles* (RUC) has been coined specifically for this project. The RUC is calculated by subtracting cyc values in all starting cycles with the cyc value in the reset event starting cycle (see Table 2.1), so that

$$RUC = cyc_{reset\ event} - cyc. \tag{2.1}$$

In effect, RUC is the concept of remaining useful life (RUL), which denotes how many cycles remain until a reset event occurs. In addition, RUC has been selected as the discrete time unit for the condition monitoring data, so that each starting cycle is enumerated by a RUC value in the countdown direction. $RUC = 0$ denotes the end of a group when the APU was "inability to start" and the faulty component was replaced through maintenance actions.

With the definition of RUC, the condition monitoring data, $info$, sn, cyc and h_{op}, that are used to record the time sequence can be replaced. The remaining condition monitoring data collected from each starting cycle can then expressed as six variables in terms of RUC,

$$\left[T_1(RUC), P_1(RUC), EGT_{peak}(RUC), N_{peak}(RUC), t_{start}(RUC), EGT_{stable}(RUC)\right], \qquad (RUC = \cdots, 2, 1, 0). \tag{2.2}$$

For simplicity, a shorten notation is also used that

$$\left[T_1, \quad P_1, \quad EGT_{peak}, \quad N_{peak}, \quad t_{start}, \quad EGT_{stable}\right]. \tag{2.3}$$

An overview of the entire dataset including all 52 groups is plotted in the Figure

2.5 with group numbers against RUC. There are a great many of points in the figure. Each point stands for an independent APU starting cycle, and corresponds to the six variables of condition monitoring data. The x axis denotes RUC value of the starting cycle, while the y axis lists the group numbers. As can be seen, theses groups have different length in terms of RUC, but all end up at the starting cycle $RUC = 0$.

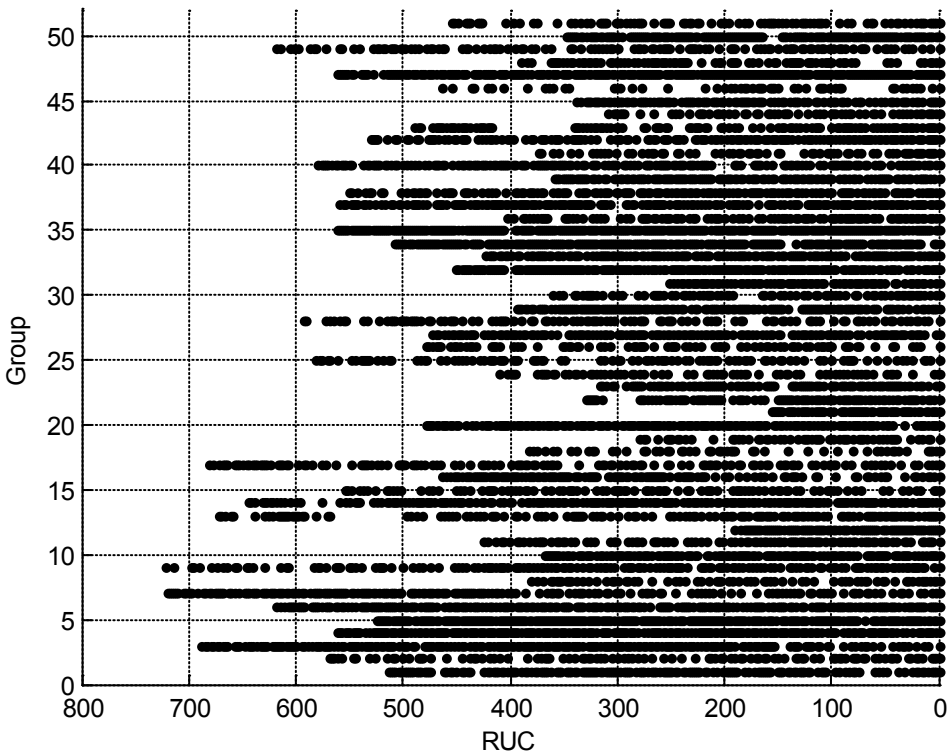


Figure 2.5 Overview of the acquired dataset $Group\ agains\ RUC$.

It is also worth mentioning that the airline maintenance program document (MPD) prescribes a breakdown maintenance strategy for the APU starter. This means that there is no periodic replacement of the starter for preventive maintenance.

Therefore, if a starter fault that led to an APU "inability to start" has existed during the 10-year period, the corresponding information will be contained within its data acquisition.

2.3 Summary

The process of data acquisition along with necessary interpretations based on APU operation information has been introduced in this chapter. Some important facts have been established, as noted below.

- In each APU starting cycle, six variables condition monitoring data have been collected: $[T_1,\ P_1,\ EGT_{peak},\ N_{peak},\ t_{start},\ EGT_{stable}]$.
- A *Reset event* has been defined as the occurrence of the APU failure "inability to start". This event indicates that, certain component has been replaced through maintenance action. However, the maintenance record does not specify which the component has been replaced.
- 52 valid *groups* have been extracted from the original dataset. Each group contains the information about one reset event.
- RUC has been selected as the discrete time unit of the condition monitoring data. RUC value also indicates the RUL.

The data acquisition step in this thesis has a great practical value due to its easy accessibility. As can be seen, all the condition monitoring data are routine sensors data collected from airline operation. Thus neither special experimental platforms nor extra

costs are required for the data acquisition. In addition, this thesis's finding can be applied not only to the aircraft A310, but also to any mainstream civil aircraft including the ERJ 190, the B737, and the A330, among others.

The dataset, however, not as intact as a designed experimental dataset or a computer simulated dataset, is interfered by the actual airline operational factors, and this can make the development of CBM framework challenging. Nevertheless, this thesis aims to tackle those challenges. Chapter 3, therefore, will address the challenge caused by the impacts of various ambient conditions (such as temperature, pressure and humidity) from actual airline operations. Chapter 4 will address the challenge caused by the fact that no sensors are installed in APU system to measure the starter performance directly. Finally Chapters 5 and 6 will address the challenge that all the collected parameters are interfered by the "noise" caused by engine operating uncertainties such as conditions of air flow, fuel-air mixture, flame propagation, etc.

Chapter 3 Data processing

The data processing performed in this chapter handles and analyzes the dataset introduced in the data acquisition step. Section 3.1 presents the correction of the raw engine operating parameters to a standard air condition in order to eliminate the effects of various ambient conditions. On the basis of the corrected results, Section 3.2 reveals that there are 3 distinctive data patterns in all groups. After that, the section lists all the possibly faulty components that could lead to the APU "inability to start" failure based on maintenance records statistics.

3.1 Gas turbine engine parameters corrections

The dataset in this thesis covers a wide range of ambient temperatures from -20° to 40° and ambient pressures relevant to the airport elevations from sea level (Vancouver) to 3557ft (Calgary). Since the ambient conditions have a significant impact on gas turbine engine performance, making the engine parameters comparable requires a correction from the actual ambient conditions to the sea level condition of *international standard atmosphere* (ISA).

The most widely applied correction method is to follow the principle of Mach number similarity. This correction ensures that the Mach number velocity triangles in engine calculated by the corrected parameters in ISA conditions are the same as the velocity triangles calculated by the raw parameters in actual ambient conditions.

Because Mach number can describe flow compressibility and velocity triangles that corresponds to the engine work and incidence losses, during this correction procedure, the parameters of engine efficiencies, temperature ratios and pressure ratios will remain unchanged. The theoretical model of EGT and N corrections based on the Mach number similarity have been presented by Volponi in [54] as follows,

$$EGT_c = \frac{EGT_{raw}}{\Theta}, \tag{3.1}$$

$$N_c = \frac{N_{raw}}{\sqrt{\Theta}}, \tag{3.2}$$

where the subscript $'c'$ stands for the corrected parameters and the Θ is the ratio of actual ambient temperature to ISA temperature, T_1/T_{ISA}.

In most practical applications, the theoretical model given by Equation (3.1) and (3.2) does not yield satisfying results. One reason for this is that the gas constant R and the ratio of specific heat γ vary due to air humidity. The other reason is the flow boundary layers vary due to the uncertainty of Reynolds number. Therefore, the quality of the correction procedure is often improved by applying empirical exponents of Θ, as shown in Equation (3.3) and (3.4) [55].

$$EGT_c = \frac{EGT_{raw}}{\Theta^{a_{EGT}}}, \tag{3.3}$$

$$N_c = \frac{N_{raw}}{\Theta^{a_N}}. \tag{3.4}$$

The empirical exponents a_{EGT} or a_N are normally determined by running a

calibrated thermodynamic computer model provided by engine manufacturers. In this thesis, the dataset has been employed instead because it also contains a considerable number of engine parameters covering a wide range of ambient conditions. The only difference is that the engine parameters simulated by the computer model are deterministic in terms of ambient conditions, whereas in the dataset, the parameters collected from actual engine operations are random variables due to engine operating uncertainties.

Taking EGT for example, the correction procedure is performed as follows. The corrected EGT corresponding to any ambient condition should be consistent with the raw EGT collected in the ISA condition, that is

$$EGT_c = \frac{EGT_{raw}}{\Theta^{a_{EGT}}} = EGT_{raw}|_{ISA}, \tag{3.5}$$

where the EGT_{raw} is a random variable as shown in Figure 3.1 (a). In this case the a_{EGT} can be determined by satisfying the least square estimation,

$$\underset{a_{EGT}}{\arg\min} \left\| \frac{EGT_{raw}}{\Theta^{a_{EGT}}} - EGT_{raw}|_{ISA} \right\|. \tag{3.6}$$

Various algorithms can be applied to solve Equation (3.6), while in this thesis the Nelder-Mead downhill simplex algorithm [56] is employed. The correction results are presented in Figure 3.1.

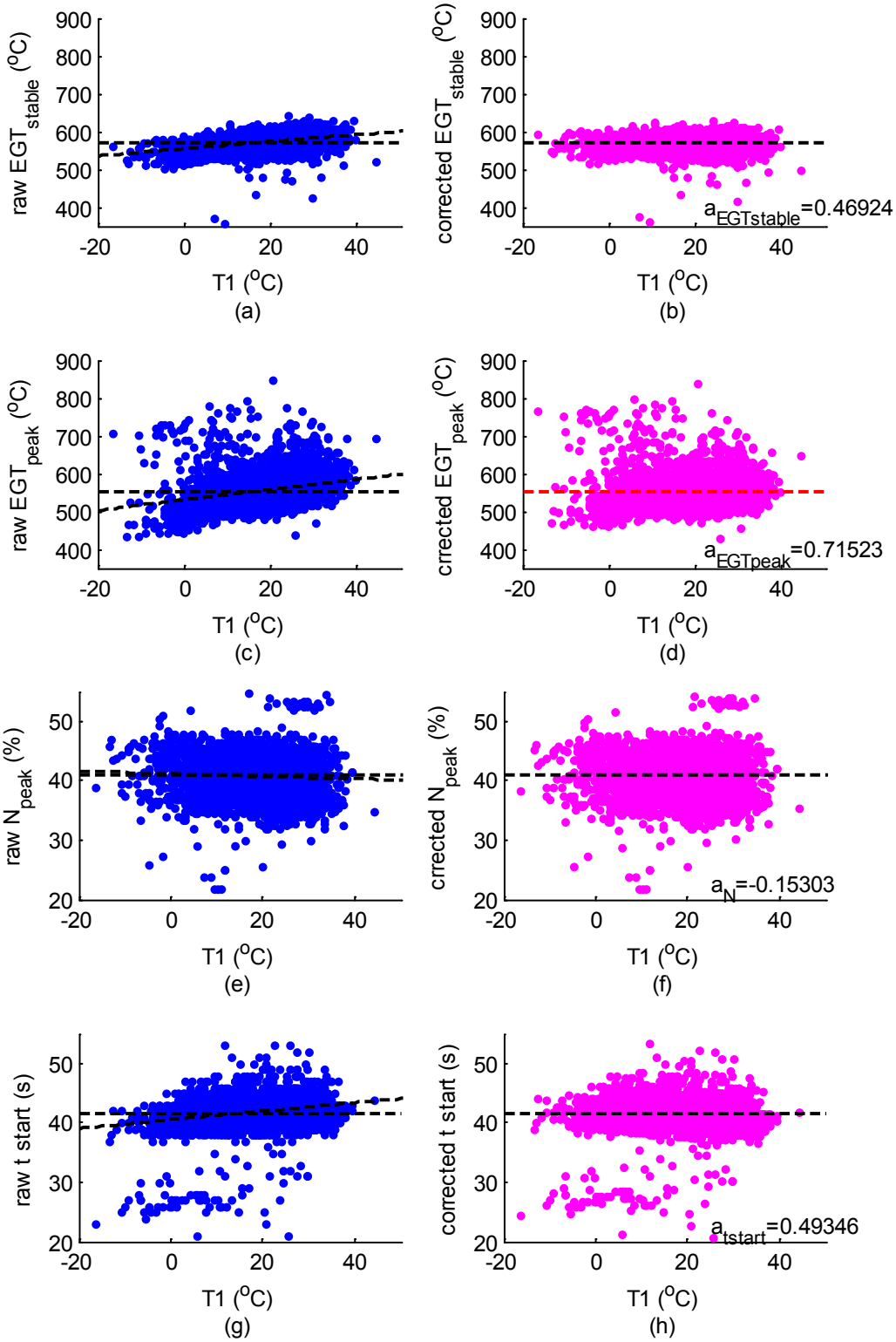


Figure 3.1 Parameters correction (a) raw EGT_{stable}, (b) corrected EGT_{stable}, (c) raw EGT_{peak}, (d) corrected EGT_{peak}, (e) raw N_{peak}, (f) corrected N_{peak}. (g) raw t_{start}, (h) corrected t_{start}.

Taking the EGT_{stable} as an example, the points in the subplot (a) represent the original EGT_{stable} measurements collected from various ambient temperature T_1, The tilted line is the least square estimation that indicates the data trend with respect to different T_1 values. The points in the subplot (b) represent the corrected EGT_{stable} values. After correction, the least square estimation of EGT_{stable_c} is invariable regardless of T_1.

The measurement and correction of EGT_{peak} are shown in the subplots (c) and (d). In contrast to EGT_{stable} that is collected after starting process with stable $100\%\ N$, the EGT_{peak} is collected during the transient starting process when EGT value peaks around $40\%\ N$. Therefore, greater engine operating uncertainties relating to such factors as air flow, fuel air mixture, flame propagation are involved in the EGT_{peak}. Two expected differences between these two variables EGT_{stable} and EGT_{peak} can be observed from the corresponding subplots. Firstly, EGT_{peak} has greater fluctuations than EGT_{stable}. Secondly, EGT_{peak} is more easily affected by T_1. Another thing worth mentioning is that EGT_{stable} is measured when the APU undertakes the bleed load of the air conditioning system and the electrical load of the aircraft. Without these loads, EGT_{stable} would be obviously lower than EGT_{peak} when collected from the same starting cycle.

As for the other two variables of condition monitoring data, N_{peak} and t_{start}, they are actually not engine thermodynamic parameters, since N_{peak} is the rotational speed at the moment of EGT_{peak} during starting and t_{start} is the time duration of

starting process. However, this correction procedure can still be applied to eliminate the effect of ambient conditions for these two parameters.

3.2 Data patterns identification

After data correction, the effects of ambient conditions are eliminated and the corrected engine parameters from different starting cycles became comparable. For simplicity, the following shortened notations will be used:

- EP: Corrected EGT_{peak}.
- NP: Corrected N_{peak}.
- ts: Corrected t_{start}.
- ES: Corrected EGT_{stable}.

The six variables of condition monitoring data in Equation (2.2) can be replaced by the following four variables in terms of RUC,

$$[\,EP, \quad NP, \quad ts, \quad ES], \tag{3.7}$$

As mentioned before, ES is collected after the starting process with $100\%\ N$. It relates to the APU working condition "after starting" rather than "during starting". The visualization of ES for all groups reveals one same data pattern. One of the groups is shown in Figure 3.2. Each point stands for a ES sample collected from the starting cycle numbered by the corresponding RUC value. The whole group ends up

at the starting cycle $RUC = 0$ when the reset event has occurred. Three separate boundaries at $RUC = 150$, 100 and 50 are selected in order to demonstrate the underlying data patterns. The whole group has then been divided into 4 zones, while the data pertaining to these zones are denoted with symbols of circle, square, triangle and star respectively. As shown in the figure, ES distributes stably around a constant value regardless of RUC countdown. This data pattern reveals that the APU has worked normally and stably after starting in all the cycles prior to the reset event.

As mentioned previously, the reset event signals an occurrence of the APU failure "inability to start" due to a certain faulty component. Although ES does not reflect APU condition during starting process, the conclusion extracted from the data patterns of this variable is important: In all the groups, if a faulty component exists that has caused "APU inability to start", this component should only affected APU working condition during starting process, not APU working condition after starting process.

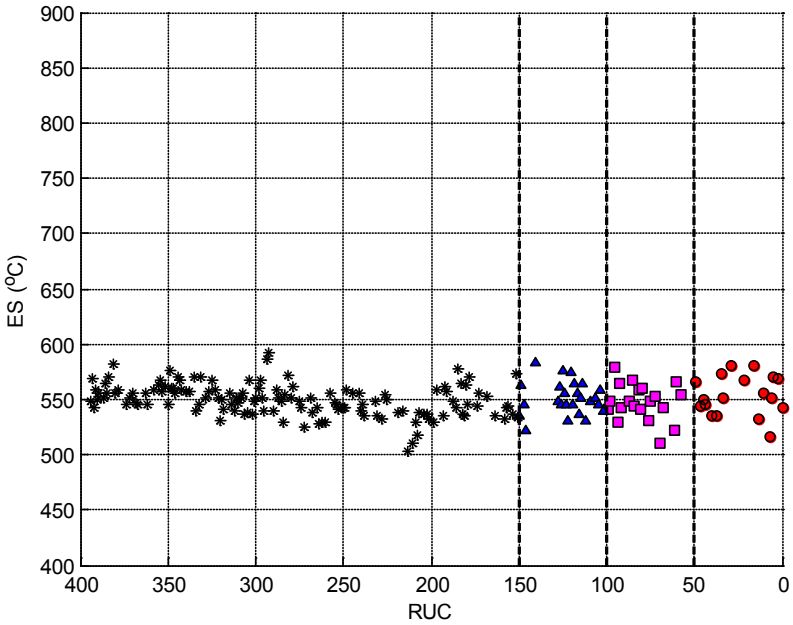


Figure 3.2 ES against RUC

Removing the ES, the remaining three variables of condition monitoring data are

$$[\,EP, \quad NP, \quad ts\,]. \tag{3.8}$$

All these three variables are related to APU starting process directly. Through the visualization of these variables, three distinctive data patterns have been identified, and all the 52 groups have been classified into three categories based on the data patterns.

The first category has 28 groups sharing the same data pattern. One group is shown in Figure 3.3 as an example. As can be seen in the subplots (a) - (c), EP, NP

and ts all maintain stable with respect to RUC. The subplot (d) shows that all points $(NP,\ EP)$ from different starting cycles cluster together without clear separations. These data patterns indicate that the starting profiles (Figure 2.4) are consistent for different cycles prior to the reset event. And there is no hint of any abnormal behaviors.

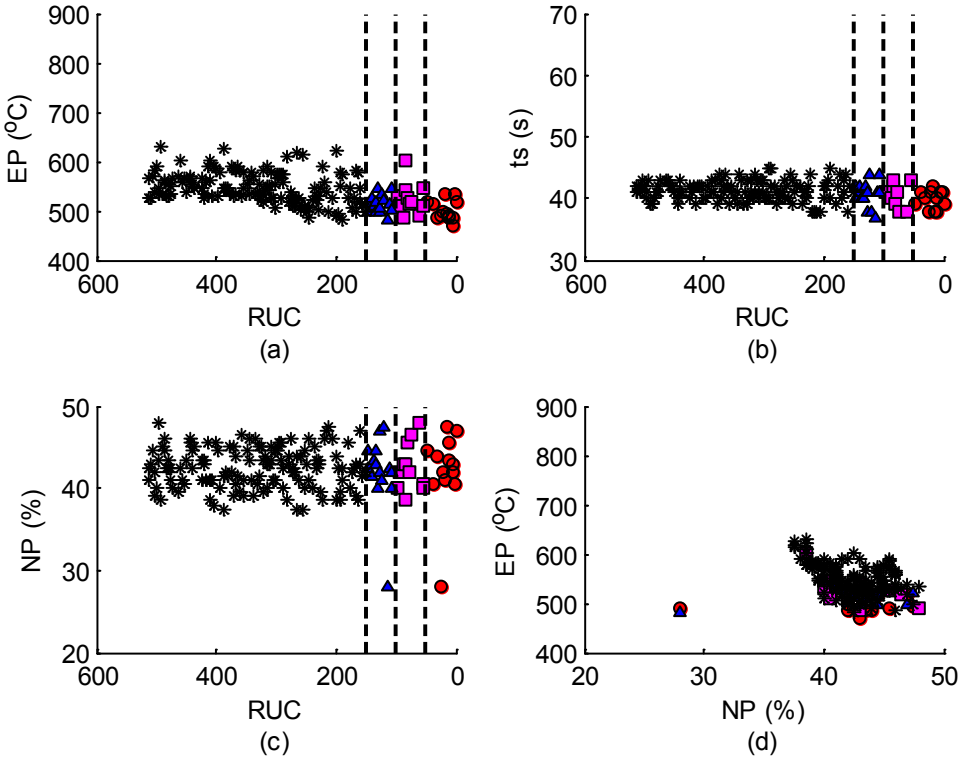


Figure 3.3 the first category data patterns (a) EP against RUC, (b) ts against RUC, (c) NP against RUC, (d) EP against NP.

The second category has 13 groups sharing another data pattern. One of the groups is shown in Figure 3.4 as an example. The developing trends of the three variables, EP, NP and ts, in terms of RUC are illustrated in the subplots (a), (b),

(c). In the zone above 150 RUC, which is far away from the reset event, their values are relatively constant. In the following zones as RUC approaches to zero, their values vary significantly in the way that EP and ts increase and NP decreases. Clearly, a certain component degrades in this zone and seriously impacts the APU starting process. In the subplot (d), the point $(NP,\ EP)$ develops toward the top left corner as RUC approaches to zero. This indicates that with a certain component degradation, the starting process gains a higher EP and an earlier associated NP.

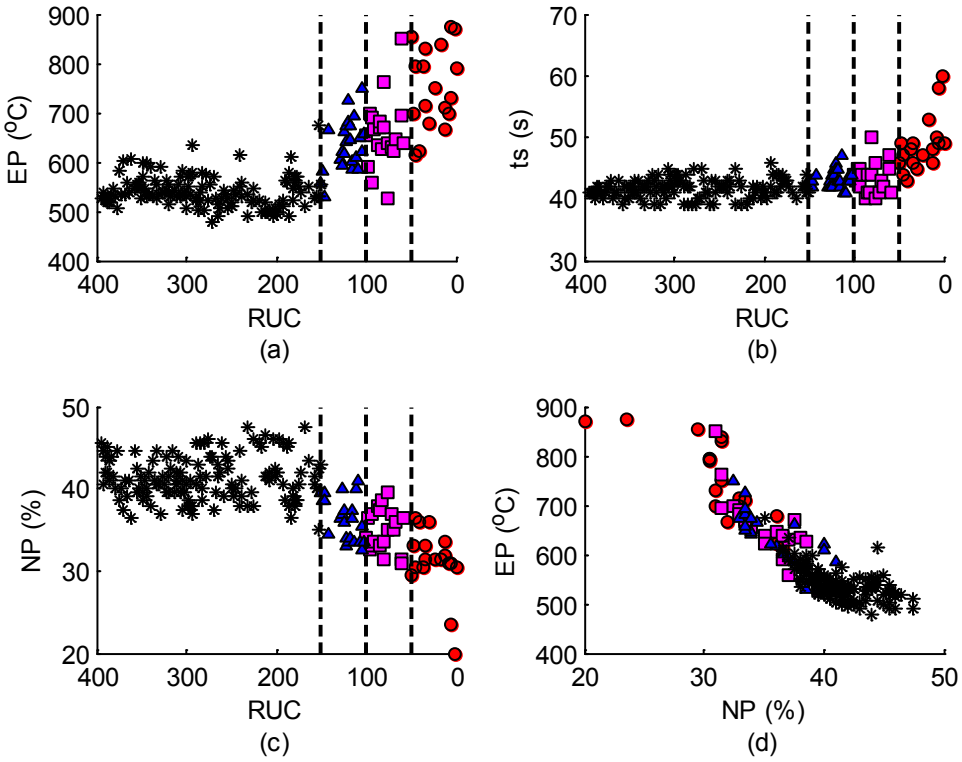


Figure 3.4 the second category data patterns (a) EP against RUC, (b) ts against RUC, (c) NP against RUC, (d) EP against NP.

The third category contains the remaining 11 groups. One of the groups is shown

in Figure 3.5 as an example. The most distinctive characteristic of this data pattern is that in the 100-0 RUC zone, two different stable values of NP exist while EP fluctuates furiously. The subplot (d) shows that the development of the points (NP, EP) is separated into two clusters. This indicates that when a certain component degrades, from different starting cycles, the same EP value occurs at a rotational speed NP that shifts back and forth between two stable values.

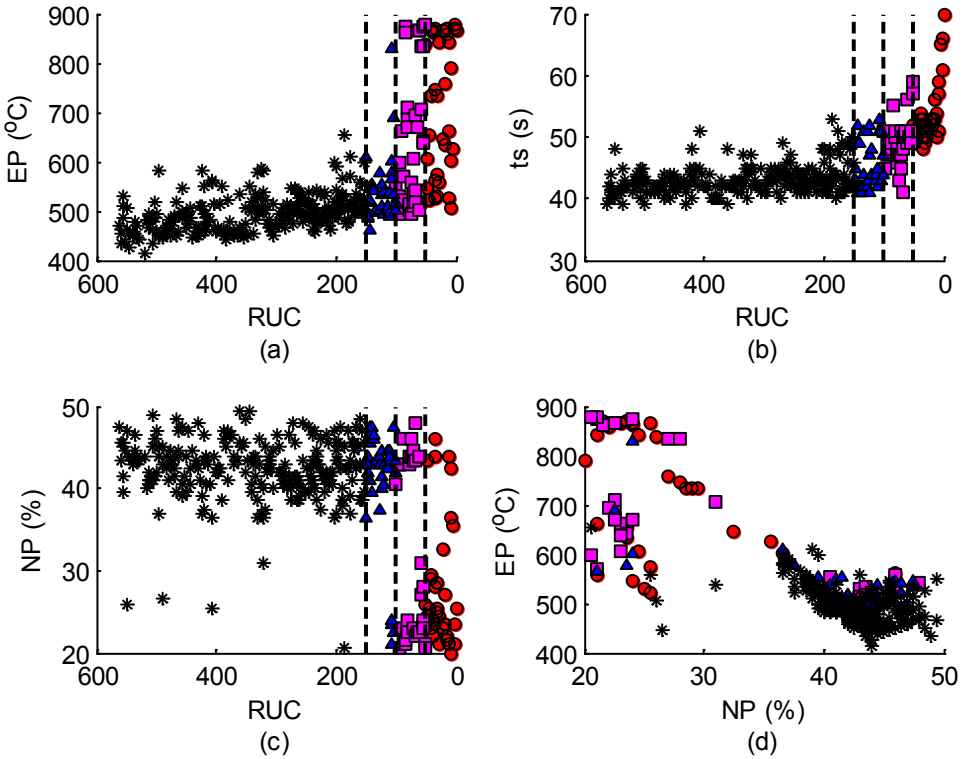


Figure 3.5 the third category data pattern (a) EP against RUC, (b) ts against RUC, (c) NP against RUC, (d) EP against NP.

These three data patterns identified above are caused by different faulty components. In paper [53], Yang *et al.* presented Failure Mode and Effects Analysis

(FMEA) of the APU. Through an investigation of Original Equipment Manufacturer (OEM), which includes maintenance record on APU systems, the paper listed all the recorded faulty components that led to the failure effect “APU inability to start”. Table 3.1 below is adapted from that paper, where the FMP denotes failures mode probability and the FR denotes the failure rate per 1 million APU operating hours. As shown in the table, the faulty components that most frequently cause APU “inability to start” are starter, igniter and fuel subsystem.

Table 3.1 FMEA of APU inability to start [53].

Component Name	FMP(%)	FR
Starter	41.4	47.61
Igniter	31.2	35.88
Fuel Control Assembly	13	14.95
Fuel Pump	5	5.75
Fuel Flow Divider	2.8	3.22
Low Oil Pressure Switch	2.2	2.53
Oil Pump Assembly	0.8	0.92
Monopole Speed Sensor	0.8	0.92
EGT Thermocouple	0.2	0.23

Two questions have been raised in this chapter. Firstly, which data pattern has

been caused by which faulty component? And secondly, if the faulty component is the APU starter, how can the three variables of condition monitoring data, EP, NP and ts, indicate the starter health condition? Both of the two questions will be answered in Chapter 4 where gas turbine performance during the starting process will be discussed.

3.3 Summary

Through the correction procedure based on the Mach number similarity, the effect of different ambient conditions from airline operations has been eliminated, and the condition monitoring data collected from different starting cycles have become comparable.

Based on the corrected exhaust gas temperature recorded when APU works stably after starting (ES), it becomes clear that all the APU's stable working conditions can be justified as normal. Therefore, only those components that have effects on APU performance during the starting process can cause the reset events. The condition monitoring data are further simplified into three variables, EP, NP and ts. The examination of these remained variables leads to the observation of three different data patterns corresponding to different component faults. In next Chapter, the engine turbine theory will be discussed with the purpose to identify the exact faulty component for different data patterns and then to establish an indicator for the starter health condition based on these three variables of condition monitoring data.

Chapter 4 Theory of gas turbine engine starting

In this chapter, the physical relationship between engine performance and APU starter health condition will be established in order to achieve two objectives. The first objective is to identify the specific faulty components that correspond to the different data patterns observed in Chapter 3. And the second objective is to establish method by which the starter health condition can be indicated by the condition monitoring data, EP, NP and ts.

Section 4.1 begins by describing gas turbine engine off-design performance when the engine is running in an *equilibrium* state (balanced state). Section 4.2 then describes the engine performance during the starting process which is a typical *transient* state (unbalanced state). Lastly, Section 4.3 describes the effect of starter degradation on engine starting performance and identifies the failure modes corresponding to different data patterns.

4.1 Off-design performance

A gas turbine engine is designed to give the required performance when stably running at the *design point*. This design point refers to a particular set of conditions including rotational speed, pressure ratio, air mass flow, etc. By contrast, when the engine is running at a condition other than its design point, the corresponding engine performance is referred as *off-design performance*. All off-design performance

calculations require the compatibilities of mass flow, work, and rotational speed between the various engine components. This section first introduces these individual engine components with the associated performance characteristics, and then demonstrates the off-design performance calculations.

The APU studied in this thesis is a single-shaft engine delivering shaft power. It contains three main components: compressor, combustor, and turbine. Figure 4.1 shows the simple diagram of this engine type. The engine's four stations are enumerated as (1) the compressor inlet, (2) the compressor outlet, (3) the turbine inlet, and (4) the turbine outlet. For simplicity, the aerodynamic parameters at different stations are labeled with the subscripts of the corresponding station numbers (e.g., P_2 denotes the pressure at the compressor outlet).

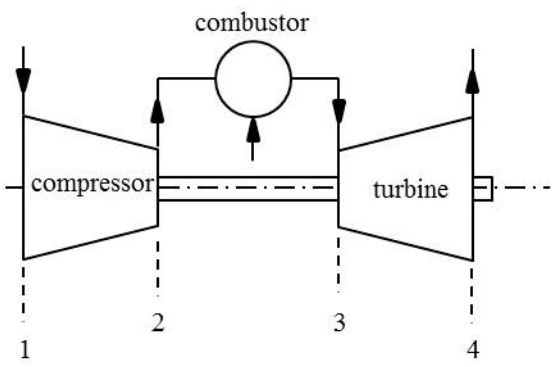


Figure 4.1 Single-shaft engine delivering shaft power.

A *compressor characteristic* (also known as a compressor map) is a series of curves that shows the compressor's performance parameters at different rotational speeds. Figure 4.2 is reproduced from the book [3]. It shows the characteristic of a

general centrifugal compressor. This characteristic plots compressor efficiency η_c and compressor pressure ratio P_2/P_1 respectively against non-dimensional air mass flow $m\sqrt{T_1}/P_1$. Each curve in the figure is called a *constant speed line* on which all points have the same non-dimensional rotational speed $N/\sqrt{T_1}$. The units of both the $m\sqrt{T_1}/P_1$ and the $N/\sqrt{T_1}$ are percentage ratios relative to the corresponding design values.

As indicated in the subplot (a), the η_c on each constant speed line varies with the $m\sqrt{T_1}/P_1$ and reaches its maximum value at a certain point. In addition, the maximum η_c values on different constant speed lines are generally consistent. In the subplot (b), the P_2/P_1 also varies with the $m\sqrt{T_1}/P_1$ on each constant speed line. The right hand extremities represent the points where choking occurs. The left hand extremities of all constant speed lines join up to form the *surge line*. And the points corresponding to the maximum η_c of all the constant speed lines join up to form the *maximum efficiency line*.

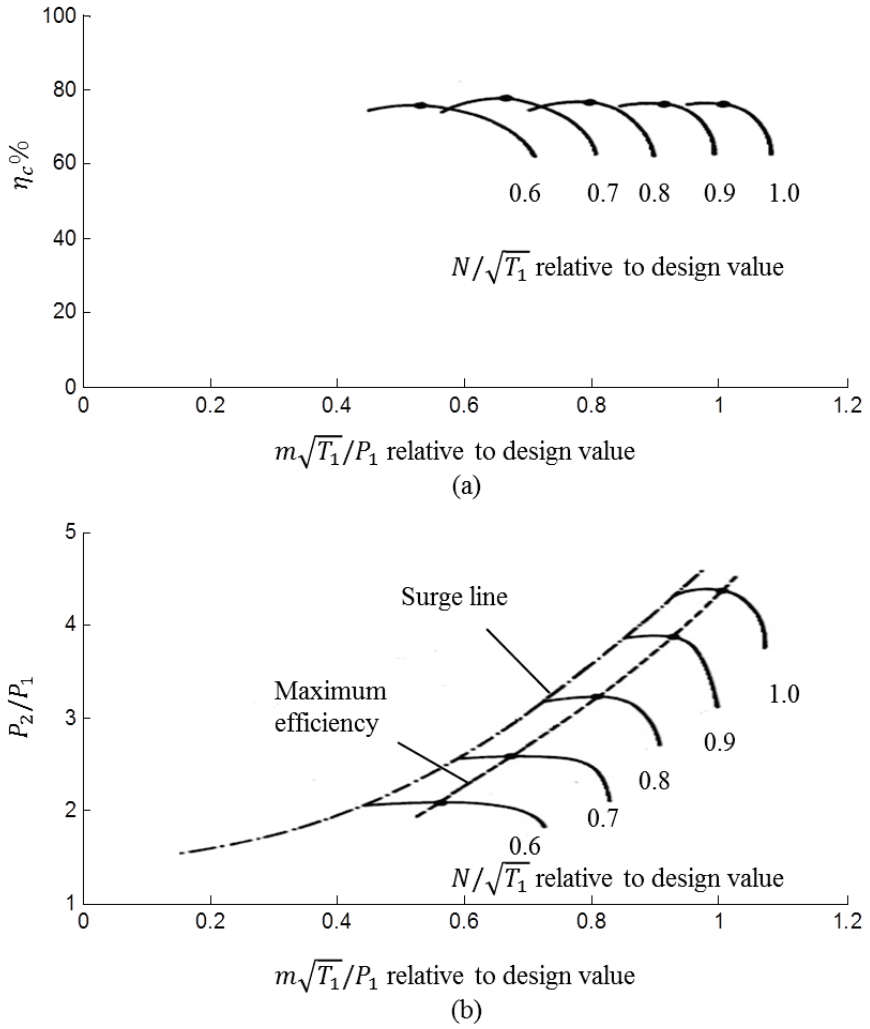


Figure 4.2 Centrifugal Compressor Characteristic [3].

Similarly, as the compressor characteristic, a general axial turbine characteristic is shown in Figure 4.3. In the subplot (b), the different constant speed lines are located closely to each other and gradually merge into a single horizontal line when the turbine pressure ratio P_3/P_4 increases. These lines merge because that a choking occurs in the turbine nozzle. When the P_3/P_4 reaches a certain threshold, the mass flow reaches its peak value and cannot further increase. In engine performance

calculation, all the constant speed lines can be approximated by the single dash curve shown in the subplot (b), so that a particular $m\sqrt{T_3}/P_3$ value corresponds to a unique P_3/P_4 value with regardless of different rotational speeds.

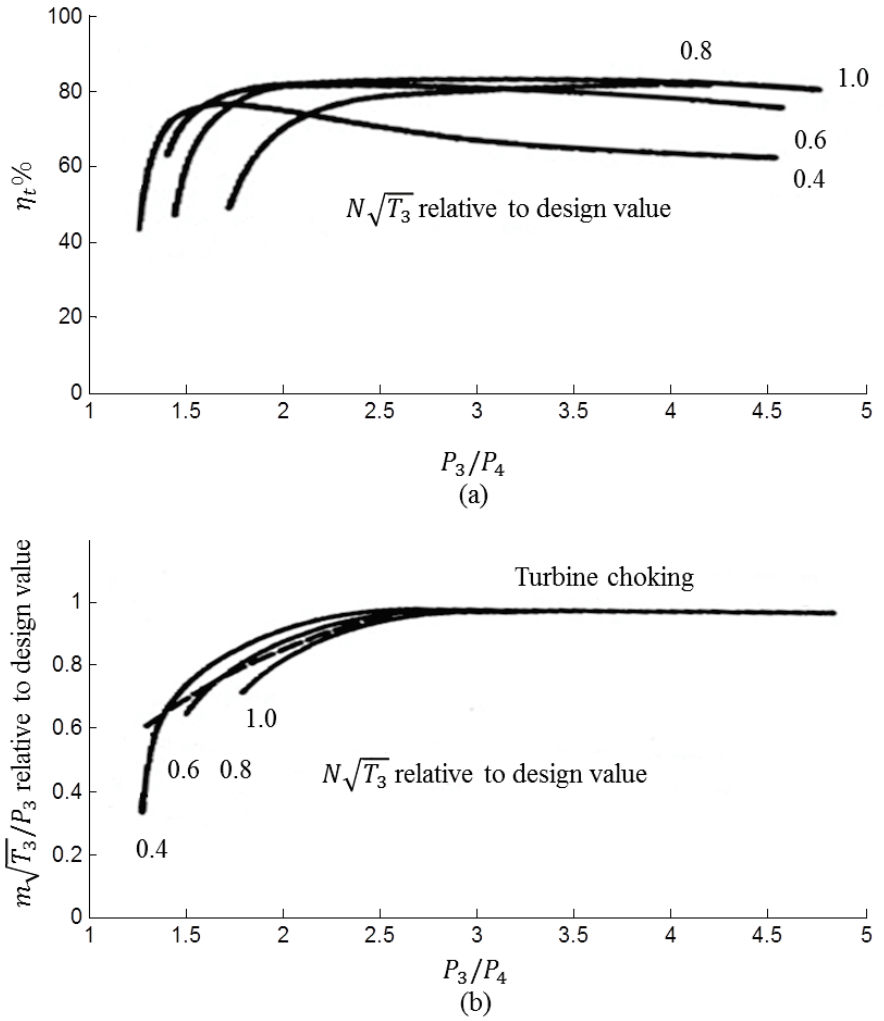


Figure 4.3 Axial turbine characteristic [3].

When an engine is running in a balanced state at an operating point within these two component characteristics mentioned above, the operating point is called *equilibrium running point*. The off-design performance at this point is calculated as

follows.

First, an arbitrary point on a constant speed line in the compressor characteristic is selected, thus the values of $m\sqrt{T_{01}}/P_{01}$, η_c, and P_{01}/P_{02} are fixed. The compatibilities of the rotational speed and air mass flow requires that

$$N_1 = N_3 = N, \tag{4.1}$$

$$m_1 = m_3 = m. \tag{4.2}$$

Equation (4.1) and (4.2) in non-dimensional forms are calculated as

$$\frac{N}{\sqrt{T_3}} = \frac{N}{\sqrt{T_1}} \times \frac{\sqrt{T_1}}{\sqrt{T_3}}, \tag{4.3}$$

$$\frac{m\sqrt{T_3}}{P_3} = \frac{m\sqrt{T_1}}{P_1} \times \frac{P_1}{P_2} \times \frac{P_2}{P_3} \times \frac{\sqrt{T_3}}{\sqrt{T_1}}. \tag{4.4}$$

Each item of Equation (4.4) is determined as follows,

- The $m\sqrt{T_1}/P_1$ and the P_1/P_2 are fixed at the selected point in the compressor characteristic.
- The P_3/P_2 is considered to be a fixed value. It is obtained from the combustion pressure loss ($\Delta P_b = P_2 - P_3$) caused by the frictional aerodynamic resistance in flame stabilizing.
- The $m\sqrt{T_3}/P_3$ is determined from the turbine characteristic with a known

value of the P_3/P_4 (the dash curve in Figure 4.3), denoted as $m\sqrt{T_3}/P_3 = f(P_3/P_4)$.

- The P_3/P_4 is equivalent to $(P_3/P_2)\cdot(P_2/P_1)$ by neglecting the inlet and exhaust pressure losses, that is $P_1 = P_4$.

Once all the values have been determined for the above parameters, the turbine inlet temperature can be obtained using Equation (4.4),

$$T_3 = \left(\frac{f\left(\frac{P_3}{P_2} \times \frac{P_2}{P_1}\right)}{\frac{m\sqrt{T_1}}{P_1} \times \frac{P_1}{P_2} \times \frac{P_2}{P_3}} \sqrt{T_1} \right)^2 . \tag{4.5}$$

Next, the turbine non-dimensional rotational speed $N/\sqrt{T_3}$ can be obtained from Equation (4.3). In addition, with the determined value of $N/\sqrt{T_3}$ and P_3/P_4, the turbine efficiency η_t can then be fixed on the turbine characteristic.

With all the items above obtained, the *compressor temperature rise* ΔT_{12} and *the turbine temperature drop* ΔT_{34} are calculated by

$$\Delta T_{12} = \frac{T_1}{\eta_c}\left[\left(\frac{P_2}{P_1}\right)^{\left(\frac{\gamma-1}{\gamma}\right)_a} - 1\right], \tag{4.6}$$

$$\Delta T_{34} = \eta_t T_3 \left[1 - \left(\frac{1}{P_3/P_4}\right)^{\left(\frac{\gamma-1}{\gamma}\right)_g}\right]. \tag{4.7}$$

The engine *net power output* $P_{w_{net}}$ is defined as the power generated by the turbine

minus the power consumed by the compressor, that is

$$P_{w_{net}} = mc_{pg}\Delta T_{34} - \frac{1}{\eta_m} mc_{pa}\Delta T_{12}. \tag{4.8}$$

In Equations (4.6), (4.7) and (4.8), the values of specific heat at the constant pressure c_p and the ratio of specific heat γ can be checked using the corresponding performance charts [57], where the subscript a stands for the ambient air while the subscript g stands for the combustion gas. The η_m is mechanical efficiency of the compressor-turbine combination by considering the aerodynamic friction, windage, and combustion loss. Finally, if an ambient temperature T_1 is chosen as a combustion reference, the fuel/air ratio f can be calculated with the known values of T_2 and T_3, that is

$$0 = (1 + f)c_{pg}(T_3 - T_1) + c_{pa}(T_1 - T_2). \tag{4.9}$$

In practical engine operation, it is also necessary to consider the external load. In order to maintain an equilibrium state, the engine net power output needs to be equivalent to the external load,

$$P_{w_{net}} - P_{load} = 0. \tag{4.10}$$

With a known external load value, the problem then becomes finding the single point

on each constant speed line of the compressor characteristic that gives the required net power output at that speed. This can be done by trial and error computer calculations.

The above procedure demonstrates that all the required parameters of off-design engine performance can be determined with a given equilibrium running point in the operation range. Repeating this procedure for a series of constant speed lines, a series of points can be obtained and joined up to form the *equilibrium running line*.

4.2 Starting process

In gas turbine operation, starting is technically the most complex phase, not only because it is a transient state that covers engine acceleration from zero to the regulated speed, but more critically, because it involves light-off interference and starter assistance. This section introduces engine starting process by describing a sequence of key phases represented by the following three figures.

Figure 4.4 profiles the power loaded onto engine spool during starting process. The upper curve stands for the engine power, with the negative values representing the *engine resistance* and the positive values representing the *engine assistance*. The lower curve stands for the starter power, referred as *starter assistance*. The starter assistance is drawn in the negative direction such that the *net unbalanced power* loaded onto the engine spool can be directly indicated by the vertical distance between two curves. Figure 4.5, which shows the same graphs displayed in Figure 2.4, is reprinted below for reading convenience. This figure records the profiles of key

APU parameters during a typical APU starting process. Figure 4.6 shows the *engine starting running line* on the compressor characteristic. Notice that this line is a transient running line rather than the equilibrium running line mentioned in the previous section, because the engine is working in an unbalanced state of acceleration on this line.

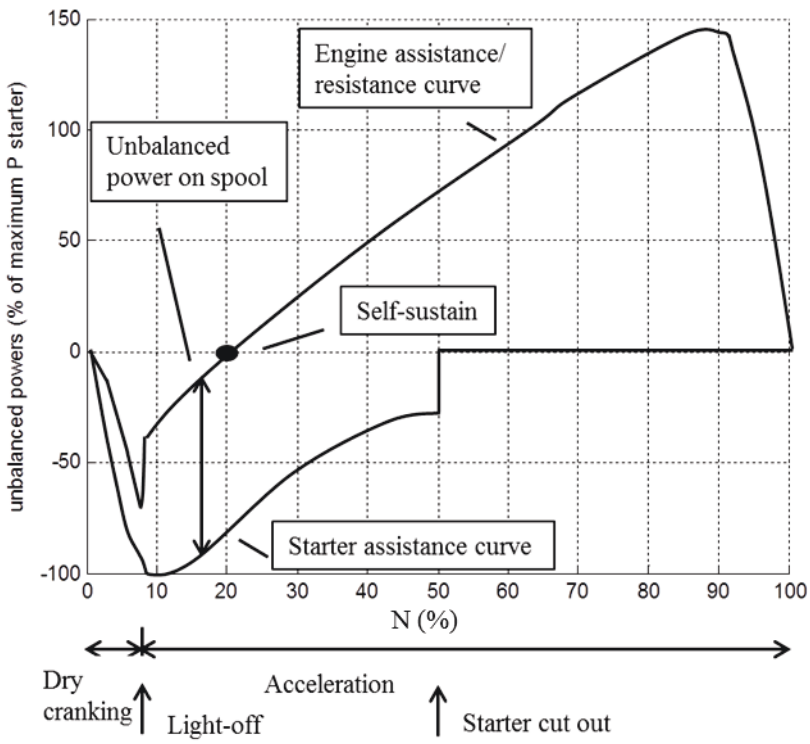


Figure 4.4 Engine resistance/assistance and starter assistance.

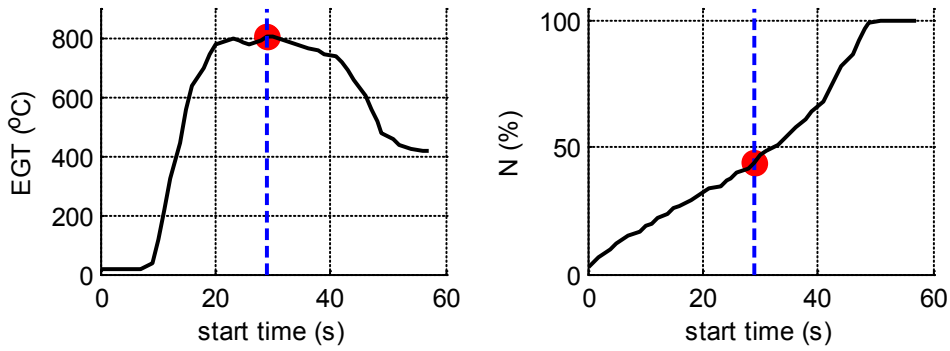


Figure 4.5 Profiles of APU starting parameters.

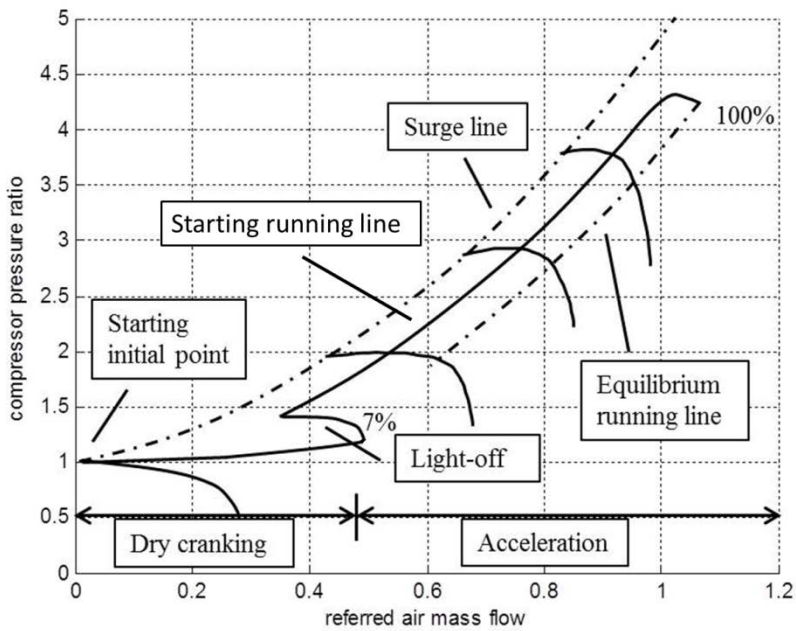


Figure 4.6 Engine starting running line.

4.2.1 Dry cranking:

The purpose of the dry cranking phase of engine starting process is to provide sufficient pressure and mass flow in the combustor for light-off.

As can be seen in Figure 4.4, during dry cranking, the starter assistance drives

the engine spool to rotate, while the engine resistance increases with cranking speed due to the flow friction and windage. The net unbalanced power is always positive in order to maintain continuous acceleration.

Figure 4.5 shows that N increases steadily due to this positive net unbalanced power. Meanwhile, since no ignition happens in this phase, EGT is equivalent to the ambient temperature.

In Figure 4.6, the engine starting running line begins from the intersection of the surge line and the zero constant speed line. Along with N increasing, the running line goes across a series of constant speed lines. In this phase, the compressor pressure ratio increases slowly as indicated by the slow rise of the running line.

4.2.2 Light-off

In the light-off phase of the engine starting process, the igniters are activated and a constant fuel flow controlled by the ECB is metered into the combustor. After the fuel is ignited by igniters, the flame propagates and stabilizes circumferentially around the combustor.

Figure 4.4 shows that a step reduction occurs in the engine resistance at the light-off moment. However, the turbine output is still less than the sum of compressor consumption and windage losses. The starter assistance is still required to continue engine spool acceleration.

In Figure 4.5, when the flame propagates around the combustor and the

combustion gas reaches the EGT thermocouples, EGT starts to rise rapidly. In this figure, the rise occurs at around $15\%\ N$.

In Figure 4.6, the engine starting running line also shows a step upwards due to the light-off. Because the engine spool inertia prevents N from sudden change, the step upwards goes along a constant speed line, which is 7% N in our case.

4.2.3 Acceleration

In the acceleration phase of engine starting process, fuel flow increases steadily along with N. In Figure 4.4, engine resistance decreases and turns into assistance after crossing the x axis. This cross point is called *self-sustain point* because the engine is able to sustain itself from this point onward. However, the starter still works for a while to facilitate a stable transition. At 50% N in our case, the starter is cut out and de-energized.

In this phase, the engine is in a transient state, so that the positive net unbalanced power continually accelerates the engine towards the target rotational speed. The transient performance can be calculated by relatively minor adjustments to the off-design performance discussed in Section 4.1.

Transient acceleration is assumed to cover a series of discrete time steps. During each time step, N is assumed to be a momentarily constant value due to the spool inertia. The fuel flow corresponding to the equilibrium state can be calculated based on Equation (4.9). In order to accelerate to next step, fuel flow is increased to generate

the extra net power $\delta P_{w_{net}}$. Considering the starter assistance, $\delta P_{w_{net}}$ can be derived from Equations (4.8) and (4.10), so that

$$\delta P_{w_{net}} = mc_{pg}\Delta T_{034} - \frac{1}{\eta_m} mc_{pa}\Delta T_{012} + P_{starter}. \tag{4.11}$$

According to the dynamic principle, the relation between torque $\delta\tau$ and power $\delta P_{w_{net}}$ is

$$\delta\tau = K_1 \frac{\delta P_{w_{net}}}{2\pi N}. \tag{4.12}$$

Since the inertia law states that the acceleration torque is equal to the product of spool inertial and angular acceleration, the following relationship is deduced:

$$\delta\tau = IK_2 \frac{dN}{dt}, \tag{4.13}$$

where K_1 and K_2 are constant values relevant to engine structure parameters. Equation (4.12) and (4.13) are then rearranged to get

$$dN = \left(\frac{K_1 \delta P_{w_{net}}}{2\pi I N K_2}\right) dt. \tag{4.14}$$

Therefore, the speed changed from one time step to the next can be described as

$$\delta N = \left(\frac{K_3 \delta P_{w_{net}}}{IN}\right) \delta t. \tag{4.15}$$

Once the above procedure is repeated for each time step, all the required parameters of engine performance in the acceleration phase can be determined.

As indicated in Figure 4.6, in this acceleration phase, the engine starting running line deviates from the equilibrium running line. It goes towards the surge line and returns back when a target speed (100% N in our case) is reached. Correspondingly, in Figure 4.5, EGT reaches its peak value shortly after the light-off moment and then gradually returns to the stable value. The reason is analyzed as below.

As mentioned before, in the transient state of acceleration, the fuel flow is higher than it is in the equilibrium state in order to provide an extra power δP_w. Meanwhile, the engine spool inertia tends to prevent N and air mass flow from increasing rapidly. Therefore the actual fuel/air ratio f is higher than it is in the equilibrium state, which causes a higher turbine inlet temperature T_3. According to Equations (4.7) and (4.8), since T_3 increases, the *turbine temperature drop* ΔT_{34} and the turbine output power must increase. According to Equation (4.6), this extra turbine power then increases the *compressor temperature rise* ΔT_{12} and the compressor pressure ratio P_2/P_1. Since the rotational speed N has hardly changed, the engine running point tends to move upwards roughly along a constant speed line. When the engine stops acceleration at the target speed, the effect just described diminishes.

4.3 Starter degradation

In this section, the effect of starter degradation on engine starting performance will be described. Then the starter health condition can be indicated by the condition monitoring data. After that, the faulty components corresponding to the three different data patterns described in Chapter 3 will be identified. Thus the objective stated in this chapter's introduction will be achieved.

4.3.1 Indication of starter health condition

Starter assistance power $P_{starter}$ is the most critical starter performance parameter required in the engine starting process. It is also the most direct numerical indication of starter health condition. When the starter is in perfect health condition, the $P_{starter}$ profile is stable and identical for different starting cycles. When the starter gradually degrades due to environmental corrosion, aging, etc., the $P_{starter}$ significantly reduces accordingly.

As mentioned in Chapter 2, there is no sensor installed in the civil aircraft that monitors and records the APU $P_{starter}$ profile directly. However, based on the gas turbine engine theory provided in this chapter, a qualitative relationship between the $P_{starter}$ and the condition monitoring data, EP, NP, and ts, can be established.

According to Equation (4.11) and (4.14), a reduced $P_{starter}$ causes a reduced

engine net power $\delta P_{w_{net}}$. As a result, the engine acceleration dN/dt is reduced and consequently, the starting time ts is extended. On the other hand, since the time span from light-off (at $7\%\,N$) to the EP (at the moment that flame propagates to the EGT sensor) is almost constant, the drop of dN/dt results in a decreased NP and a decreased air mass flow in the combustor. Since the fuel supply at light-off is a fixed value controlled by the ECB, the fuel/air ratio f thus increases. This, in turn, eventually causes EP to increase.

Figure 4.7 illustrates the relationship between the $P_{starter}$ and the profile of EGT against N. As can be seen, in the starting cycle with a reduced $P_{starter}$, EP is higher and NP is lower than these values would be in a normal starting cycle. The point (NP, EP) corresponding to a reduced $P_{starter}$ thus moves toward the top left. The t_{start} corresponding to a reduced $P_{starter}$ is higher than the t_{start} of a normal starting cycle, although it is not shown in this figure.

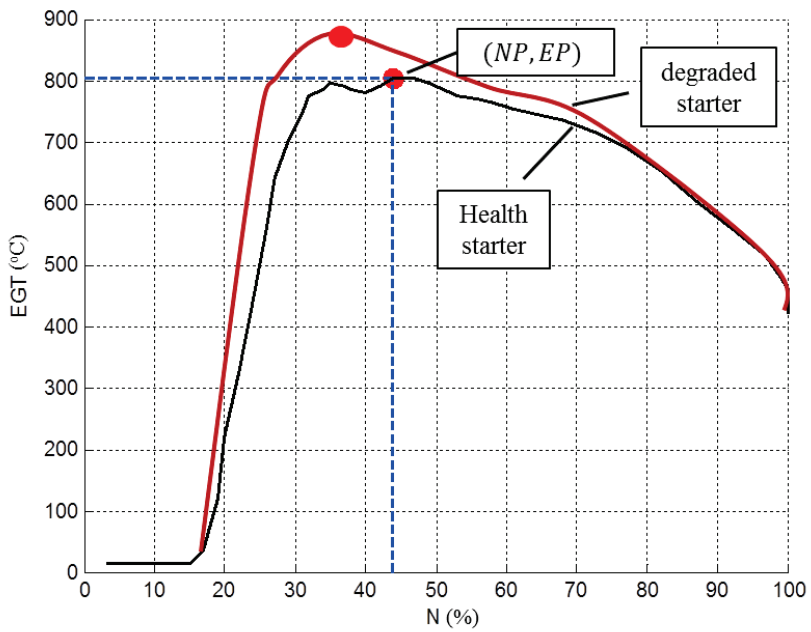


Figure 4.7 EGT against N with reduced $P_{starter}$

Based on the above reasoning, the starter health condition is numerically indicated by the starter assistance power $P_{starter}$. The $P_{starter}$ can be numerically indicated by the condition monitoring data, EP, NP, and ts. That is, the normal values of EP, NP, ts together indicates a normal $P_{starter}$ and a healthy starter, while an increased EP, decreased NP and extended ts indicate a reduced $P_{starter}$ and a degraded starter. Theoretically, a particular starter health condition would corresponds to a unique set of the condition monitoring data values, and the worse the starter degrades, the further these data deviate from normal values. However, in practical engine operation, the condition monitoring data will also be affected by uncertain conditions such as air flow, fuel-air mixture and flame propagation. These conditions cannot be exactly the same for different starting cycles. Therefore the condition monitoring data are statistical variables rather than a deterministic functions

with respect to the $P_{starter}$. These statistical properties will be explored in Chapter 5.

4.3.2 Faulty component identification

Based on all the knowledge introduced earlier in this chapter, the faulty components for the three distinct data patterns described in Section 3.2 can now be identified.

In the first category, the 28 groups corresponding to the data pattern shown in Figure 3.3 have experienced either the igniter failure or no component faultiness at all. Igniters are used to ignite fuel in the primary zone of combustor. Once an igniter failure occurs, the only phenomena that can be observed are the failure of light-off and the subsequent APU "inability to start" failure. Therefore, there is no condition monitoring data generated from the failed starting cycle, and all the data from the previous cycles behave as normal. The other possible situation that this data pattern might reflect is that all the APU components are normal but one of them is replaced due to some maintenance plans.

In the second category, the 13 groups corresponding to the data pattern shown in Figure 3.4 have experienced the APU starter degradation. As mentioned previously, a reduced $P_{starter}$ is indicated by the fact that EP increases, NP decreases and ts extends. Based on this analysis, Figure 3.4 shows that $P_{starter}$ maintains a constant condition at the beginning of the group, but decreases significantly after passing a boundary line that exists between $150 - 100\ RUC$.

In the third category, the 11 groups corresponding to the data pattern shown in Figure 3.5 have experienced a failure in the APU fuel system. As can be seen in the subplots (c) and (d), in the $100 - 0\ RUC$ zone, a particular EP value corresponds to two distinctive NP values. This shows that the fuel supply varies back and forth between two values in different stating cycles.

As discussed above, only the 13 groups in the second category contain information about APU starter degradation. Only these groups of data, therefore, will be analyzed in the following chapters for APU starter fault diagnostics and failure prognostics.

4.4 Summary

In this chapter, the gas turbine theory of off-design performance and starting process has been introduced. Accordingly, the effect of starter degradation on engine performance has been analyzed. The condition monitoring data, EP, NP and ts, have been applied to indicate the starter assistance power $P_{starter}$. Furthermore, the 13 groups of data in the second categories initially described in Chapter 3 have been identified as a useful and informative source for estimating the APU starter degradation.

Chapter 5 Diagnostics with online classifier

Diagnostics is a posterior data analysis that utilizes historical and current condition monitoring data to detect faults of targeted equipment. In this chapter, the statistical properties of the condition monitoring data will be described in Section 5.1. Then the methodologies of moving average and autocorrelation will be briefly introduced in Section 5.2. Last, an online classifier based on autocorrelation to detect the APU starter degradation will be developed in Section 5.3.

5.1 Statistical property evaluation

In the previous chapters, totally 13 data groups derived from the 10-year operation records of 35 A310 aircrafts were found to contain useful information about APU starter performance degradation. Also the relationship between the starter assistance power, $P_{starter}$, and the three condition monitoring variables, $EP(RUC)$, $NP(RUC)$, and $ts(RUC)$, have been investigated, which concludes that a decreased $P_{starter}$ leads to an increased EP, an decreased NP, and an increased ts. One group of these variables is selected as an example and plotted in Figure 5.1. The statistical properties of these variables will be analyzed in this section. For simplicity, the analysis will refer only to EP, but they apply to NP and ts as well.

In theory, a particular EP value should correspond to a unique $P_{starter}$ condition. However, in practice, the EP value is also affected by engine operating

uncertainties such as the conditions of air flow, fuel air mixture, and flame propagation. Therefore, the $EP(RUC)$ collected from an entire group exhibits as a *random process* [58], where the RUC is the time unit in a countdown direction. And each RUC value corresponds to an independent APU starting cycle.

According to the information above, the random process $EP(RUC)$ is considered as a combination of two sub-processes. The first sub-process is referred as *starter signal*. It is a deterministic process with respect to the $P_{starter}$ that indicates the starter health condition. The second sub-process is referred as *noise signal*. It is a random process caused by the engine operating uncertainties.

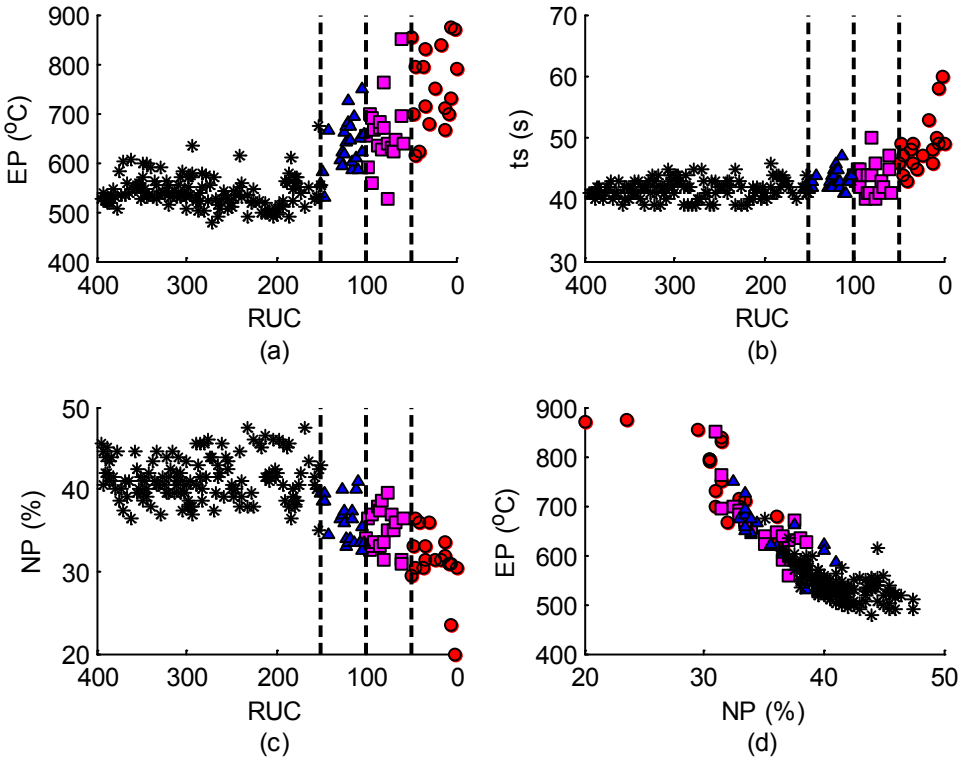


Figure 5.1 Starter degradation data pattern: (a) $EP(RUC)$, (b) $ts(RUC)$, (c) $NP(RUC)$, (d) $EP(RUC)$ against $NP(RUC)$.

In Figure 5.1(a), the random process $EP(RUC)$ from the first half and the second half of the example group exhibit two distinct distributions. In the first half, the EP values from different RUC starting cycles are relatively stable. This data pattern indicates that the starter is healthy with the stable and identical $P_{starter}$ for different starting cycles. Hence the first half is named *normal phase*. In the second half, the EP values increase considerably as RUC decreases. This data pattern indicates that the starter continually degrades with the $P_{starter}$ decreasing from one starting cycle to another. Hence the second half is named as *degraded phase*. The boundary between the normal and degraded phases appears to be located between 150 and 100 RUC. The exact location of this boundary will be determined in the later sections of this chapter. However, at this moment, it is safe to say that the EP samples collected before $150\ RUL$ belong to the normal phase.

In order to describe the underlying distributions of the normal and degraded phases, a *kernel smoother* is applied. Kernel smoother is a statistical technique for estimating probability density function (pdf) by using random experimental samples when no parametric model for this pdf is known [59]. As Figure 5.1(a) shows, three vertical lines have been drawn at 150, 100, and 50 RUC to divide the group into 4 zones. The first zone (> 150 RUC) includes EP samples from the normal phase, while the other three zones (150-100, 100-50 and 50-0 RUC) include EP samples showing increasingly serious levels of degradation. The results of the kernel smoother applied on the above zones are shown in Figure 5.2.

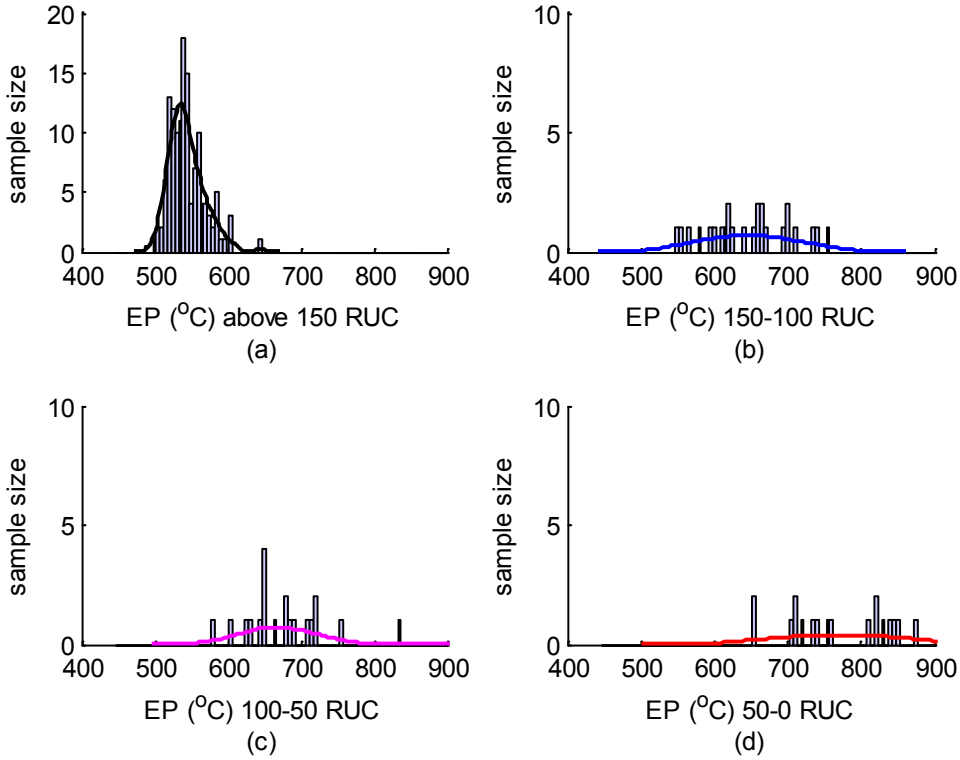


Figure 5.2 Kernel smoother of EP samples from zones: (a) above 150 RUC, (b) 150-100 RUC, (c) 100-50 RUC, (d) 50-0 RUC.

Figure 5.2 (a) shows the kernel smoother applied on EP samples in the first zone (> 150 RUC). The result indicates that the pdf of the normal phase behaves as a stationary Gaussian distribution. This is an anticipated result, since the central limit theorem states that the physical quantities that are expected to be the sum of many independent processes often have a distribution very close to Gaussian [60]. This Gaussian distribution of the normal phase is denoted as $\mathcal{N}(\mu_{nor}, \sigma_{nor}^2)$ where the μ_{nor} and σ_{nor} are the mean and standard deviation respectively. Since the $EP(RUC)$ is the sum of the starter signal and the noise signal, in the normal phase,

the underlying distribution, $\mathcal{N}(\mu_{nor}, \sigma_{nor}^2)$, can be separated as a constant value μ_{nor} that indicates the starter signal and a zero mean stationary Gaussian $\mathcal{N}(0, \sigma_{nor}^2)$ that is the noise signal.

Figure 5.2 (b), (c) and (d) show the kernel smoothers applied to the other three zones ($150 - 100$, $100 - 50$ and $50 - 0$ RUC). From one zone to the next, the mean value of EP increases significantly. This indicates a continuous starter degradation in the degraded phase. At the same time, from one zone to the next, the standard deviation increases significantly. This happens firstly because less mass flow and more fierce combustion give rise to greater engine operating uncertainties. But this also happens because of the shifting of starter signal. Obviously, the distribution of the degraded phase is a non-stationary distribution.

As for the other two condition monitoring variables, $NP(RUC)$ and $ts(RUC)$, they are also considered random processes and exhibit similar properties to those found in $EP(RUC)$. In fact, the *statistical correlation* shows that the three random processes have high statistical dependence with each other. In the example group shown above, their statistical correlations are:

$$corr(EP, NP) = -0.8881,$$

$$corr(ts, NP) = -0.8277,$$

$$corr(ts, EP) = 0.7815.$$

For instance, Figure 5.2 (d) shows that EP and NP have an approximately linear relationship.

With the statistical properties discussed above, two methodologies, *moving*

average and *autocorrelation*, will be involved for the online data analysis. The methodologies are traditional but effective for identifying the starter signal and describing the data distributions. They will be briefly introduced in the next section.

5.2 Methodology

5.2.1 Moving average

Moving average is a type of finite impulse response (FIR) filter used to analyze a dataset by creating a series of averages of different subsets. It is commonly used with time series data to smooth out short-term fluctuations and highlight longer-term trends. The simple moving average (SMA) for a random process $X(t)$ is the unweighted mean of the previous ξ date points.

$$SMA_t = \frac{X(t) + X(t-1) + \cdots + X(t-\xi-1)}{n} \tag{5.1}$$

5.2.2 Theoretical autocorrelation

Autocorrelation is an important descriptor of random processes that is widely used in machinery diagnostics [61], [62]. It can identify the distribution features that are indistinguishable by mean and standard deviation. The *theoretical autocorrelation*

function for a random process $X(t)$ is defined as

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)], \tag{5.2}$$

where t_1 and t_2 are arbitrary sampling times. $R_X(t_1, t_2)$ indicates how well X correlate with itself at two different times. If $X(t)$ is a stationary process so that its pdf is invariant with time, the autocorrelation depends only on the *time difference* variable $\tau = t_2 - t_1$. Therefore, $R_X(t_1, t_2)$ is a function only of τ,

$$R_X(\tau) = E[X(t)X(t+\tau)]. \tag{5.3}$$

If the mean value of $X(t)$ is subtracted out, Equation (5.3) then becomes the *theoretical autocovariance* function, so that

$$R'_X(\tau) = E\{[X(t) - \mu_X][X(t+\tau) - \mu_X]\}. \tag{5.4}$$

Clearly, $R_X(0)$ is the mean square of $X(t)$ while $R'_X(0)$ is the variance of $X(t)$. In the case that $X(t)$ is a zero mean random process, $R_X(\tau)$ and $R'_X(\tau)$ are then identical.

From the autocorrelation definition, the quality of $R_X(\tau)$ depends on how rapidly $X(t)$ changes with respect to the time difference τ. Therefore, $R_X(\tau)$ contains the frequency information of $X(t)$. There is the theorem known as

Wiener-Khinchin to describe the relationship between the autocorrelation and the *power spectral density function* [63]. It is expressed as follows:

$$S_X(j\omega) = \mathcal{F}[R_X(\tau)] = \int_{-\infty}^{\infty} R_X(\tau) e^{-j\omega\tau} d\tau. \tag{5.5}$$

Both $R_X(\tau)$ and $S_X(j\omega)$ are important functions for describing the underlying distribution of random processes. Different random process will also show different patterns of $R_X(\tau)$ and $S_X(j\omega)$.

As shown in Figure 5.3, for a *Gaussian white noise process* $X_1(t)$, the corresponding theoretical autocorrelation is an impulse function and the power spectral density function is a constant value, such that

$$\begin{aligned} R_{X_1}(\tau) &= \sigma_{X_1}^2 \delta(\tau), \\ S_{X_1}(j\omega) &= \sigma_{X_1}^2, \end{aligned} \tag{5.6}$$

where $\sigma_{X_1}^2$ is the variance of $X_1(t)$. Equation (5.6) indicates that $X_1(t)$ has no self-correlation at any two different times, and covers equally on $(0 - 2\pi)$ ranges in frequency domain.

For another random process, taking the *first order Gauss-Markov process* $X_2(t)$ for instance, the theoretical autocorrelation and power spectral density are,

$$R_{X_2}(\tau) = \sigma_{X_2}^2 e^{-\beta_{X_2}\tau},$$

$$S_{X_2}(j\omega) = \frac{2\sigma_{X_2}^2}{\omega^2 + \beta_{X_2}^2}, \tag{5.7}$$

where $\sigma_{X_2}^2$ is the variance and $1/\beta_{X_2}$ is the estimated time constant [64].

5.2.3 Experimental autocorrelation

The determination of autocorrelation from experimental sample is a common engineering problem. Suppose $X_T(t)$ is a finite-length experimental sample of the random process $X(t)$, where the subscript T is the time length, the autocorrelation determined from $X_T(t)$ is then referred as *experimental autocorrelation* $V_X(\tau)$. It is calculated from

$$V_X(\tau) = [time\ avg.\ of\ X_T(t)X_T(t-\tau)]$$

$$= \frac{1}{T-\tau}\int_0^{T-\tau} X_T(t)X_T(t+\tau)\,dt. \tag{5.8}$$

The experimental autocorrelation $V_X(\tau)$ does not agree perfectly with the known theoretical autocorrelation $R_X(\tau)$. However, the $V_X(\tau)$ will represent the general form of the $R_X(\tau)$ in a certain accuracy bound. This accuracy bound of $V_X(\tau)$ for a stationary Gaussian process has been proved in the literature [65] that the variance of $V_X(\tau)$ satisfies the inequality,

$$Var\,[V_X(\tau)] < \frac{4}{T}\int_0^{\infty} R_X^2(\tau)d\tau, \tag{5.9}$$

where T is the time length of the experimental record and $R_X(\tau)$ is the theoretical autocorrelation function of the Gaussian process under consideration.

Figure 5.3 show the simulated autocorrelations and power spectral density functions for two random processes. $X_1(t)$ is the Gaussian white noise process and $X_2(t)$ is the first order Gauss-Markov process. They are simulated with the same mean and variance values. As can be seen, the theoretical autocorrelations of these two random processes are distinctive. $R_{X_1}(\tau)$ is an impulse function with $R_{X_1}(1) = R_{X_1}(2) = 0$, whereas $R_{X_2}(\tau)$ is an exponential function with $R_{X_2}(1) \gg 0$ and $R_{X_2}(2) \gg 0$. The experimental autocorrelations are not in perfect agreement with the corresponding theoretical autocorrelations, but fluctuate in the certain bounds.

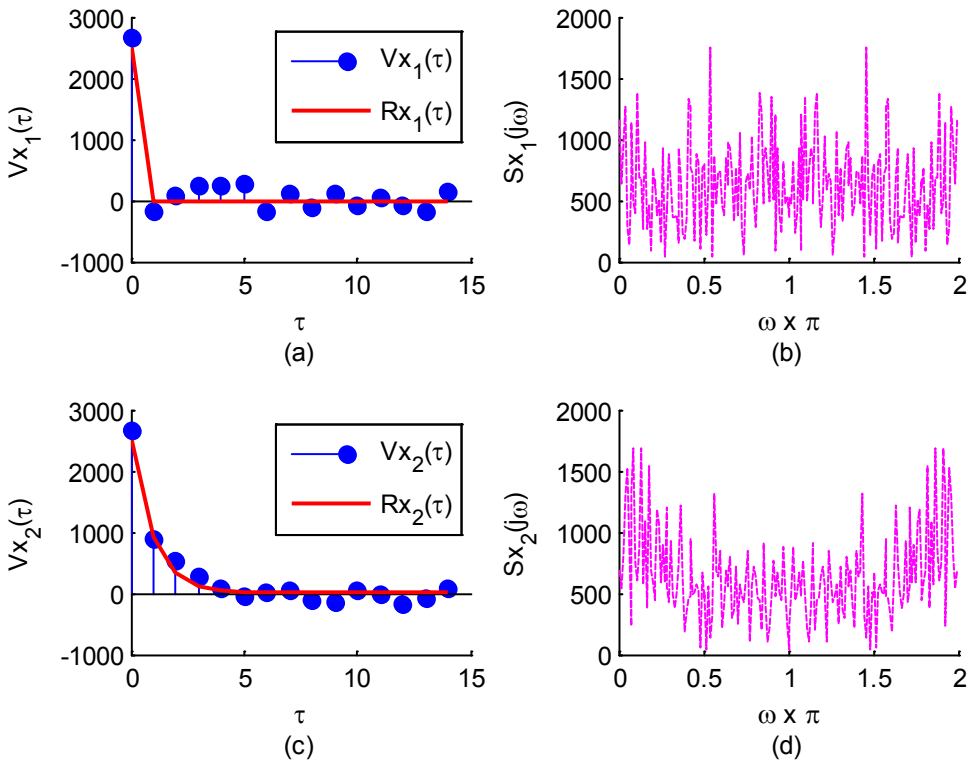


Figure 5.3 (a) autocorrelation of the Gaussian white noise process, (b) power spectral density of the Gaussian white noise process, (c) autocorrelation of the first order Gauss-Markov process, (d) power spectral density of the first order Gauss-Markov process.

5.3 Online classifier

In this section, an online classifier will be designed to detect the starter degradation by determining the boundary line between the normal and degraded phases. The procedure is shown as follows.

For each RUC, the previous 30 samples are drawn to form a subset X_{RUC} that

$$X_{RUC} = [EP(RUC + 29), EP(RUC + 28), \dots, EP(RUC - 1), EP(RUC)]. \quad (5.10)$$

With the known values of the above 30 samples, the mean $\mu_{X_{RUC}}$ and the standard deviation $\sigma_{X_{RUC}}$ of X_{RUC} can be calculated. The moving average is then obtained by joining the $\mu_{X_{RUC}}$ for a series of RUC. In Figure 5.4 (a), $\mu_{X_{RUC}}$ with bounds of $\pm 2\sigma_{X_{RUC}}$ are plotted on the $EP(RUC)$ samples.

In addition, for each X_{RUC}, the inside 30 samples are subtracted by the $\mu_{X_{RUC}}$, so that X_{RUC} is transformed into a zero mean subset,

$$X'_{RUC} = \left[EP(RUC + 29) - \mu_{X_{RUC}}, EP(RUC + 28) - \mu_{X_{RUC}}, \dots, EP(RUC) - \mu_{X_{RUC}}\right]. \quad (5.11)$$

With the new subset X'_{RUC}, the experimental autocorrelation $V_{X'_{RUC}}(\tau)$ at time difference variable $\tau = 0,1,2$, can then be calculated from (5.8). In Figure 5.4 (b), three curves, referred as *moving experimental autocorrelations*, are obtained by joining the absolute values of the $V_{X'_{RUC}}(0)$, $V_{X'_{RUC}}(1)$, and $V_{X'_{RUC}}(2)$ for a series of RUC.

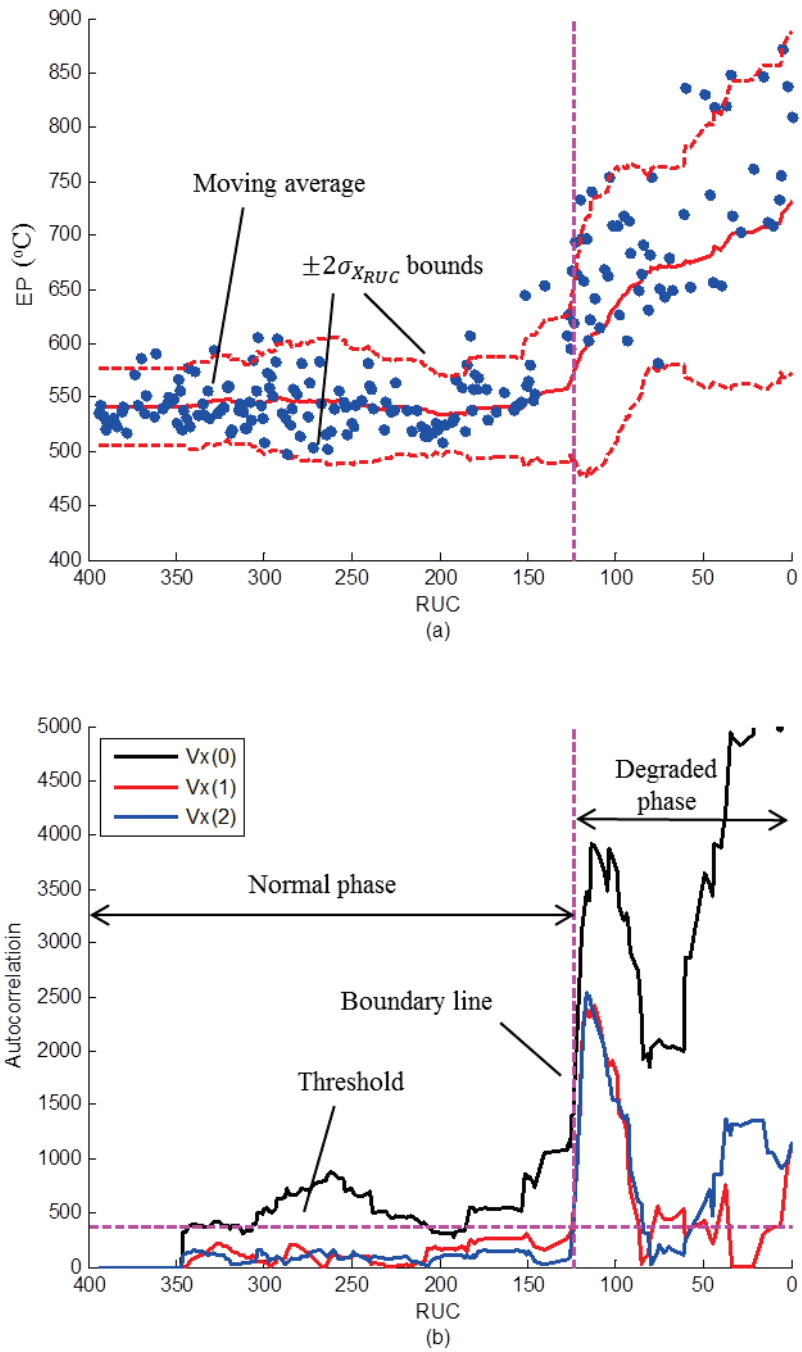


Figure 5.4 EP (a) samples, moving average, and moving standard deviation bounds (b) absolute values of moving experimental autocorrelations, the threshold, and the boundary line.

Figure 5.4 (a) shows that the moving average $\mu_{X_{RUC}}$ and the moving standard deviation $\sigma_{X_{RUC}}$ are relatively stable in the normal phase, but increase dramatically in

the degraded phase. This result reaffirms the conclusion made in Section 5.1. The conclusion states that the random process EP is a combination of the starter signal and the noise signal. In normal phase, EP satisfies the stationary Gaussian $\mathcal{N}(\mu_{nor}, \sigma_{nor}^2)$. The starter signal is indicated by the constant μ_{nor} and the noise signal satisfies a zero-mean stationary Gaussian $\mathcal{N}(0, \sigma_{nor}^2)$.

Furthermore, Figure 5.4 (b) shows that the $V_{X'_{RUC}}(1)$, and $V_{X'_{RUC}}(2)$ vary slightly around zero in the normal phase, but increase dramatically in the degraded phase. This result indicates that the noise signal is a Gaussian white noise process in the normal phase, but is a different random process in degradation phase. The mission of identifying the exact boundary line between the normal and degraded phases then becomes identifying the specific RUC where the associated subset X'_{RUC} begins to deviate from the Gaussian white noise process in the normal phase to the random process in the degraded phase.

According to Equation (5.9), in the normal phase the variance of $V_{X'_{RUC}}(\tau)$ should satisfy the inequality

$$Var\left[V_{X'_{RUC}}(\tau)\right] < \frac{4}{30}\int_0^{\infty} R_{X'_{RUC}}^2(\tau)d\tau, \tag{5.12}$$

where $R_{X'_{RUC}}(\tau) = \sigma_{nor}^2\delta(\tau)$ is the theoretical autocorrelation of the Gaussian white noise in the normal phase. The $\sigma_{nor} \approx 22.18$ can be determined from a sufficient number of samples in the zone above 150 RUC. Equation (5.12) is calculated to yield

$$\sigma_{V_{X'_{RUC}}} < \sqrt{\left(\frac{4}{30}\sigma_{nor}^4\right)} = 189.8. \tag{5.13}$$

Statistically, more than 95% of the $V_{X'_{RUC}}(\tau)$ should fall into the $\pm 2\sigma_{V_{X'_{RUC}}}$ band. In other words, once the $V_{X'_{RUC}}(1)$ or $V_{X'_{RUC}}(2)$ exceed the band, the corresponding set X'_{RUC} most probably belongs to a distribution other than $\mathcal{N}(0, \sigma_{nor}^2)$, hence the corresponding RUC can be judged to be in the degraded phase rather than in the normal phase.

Based on the above criterion, a threshold is set to be $2\sigma_{V_{X'_{RUC}}} = 379.6$ as the horizontal line shown in Figure 5.4 (b). The RUC where $V_{X'_{RUC}}(1)$ or $V_{X'_{RUC}}(2)$ exceeds this threshold is determined as the boundary line between the normal and degraded phases as the vertical line shown in Figure 5.4 (b). So far, the online classifier is achieved.

Applying the classifier to the other two random processes, NP and ts, the results are shown in Figure 5.5 and Figure 5.6. Comparing these figures, it is found that EP is most sensitive to starter degradation and can effectively determine the boundary line. NP shows almost the same result as EP. However, ts has quite slow response and cannot identify the degraded phase in time. Based on these observations, the boundary line determined by EP is selected. In addition, the ts is removed from the condition monitoring data, only the remaining two variables, EP and N, will be kept in the following analysis.

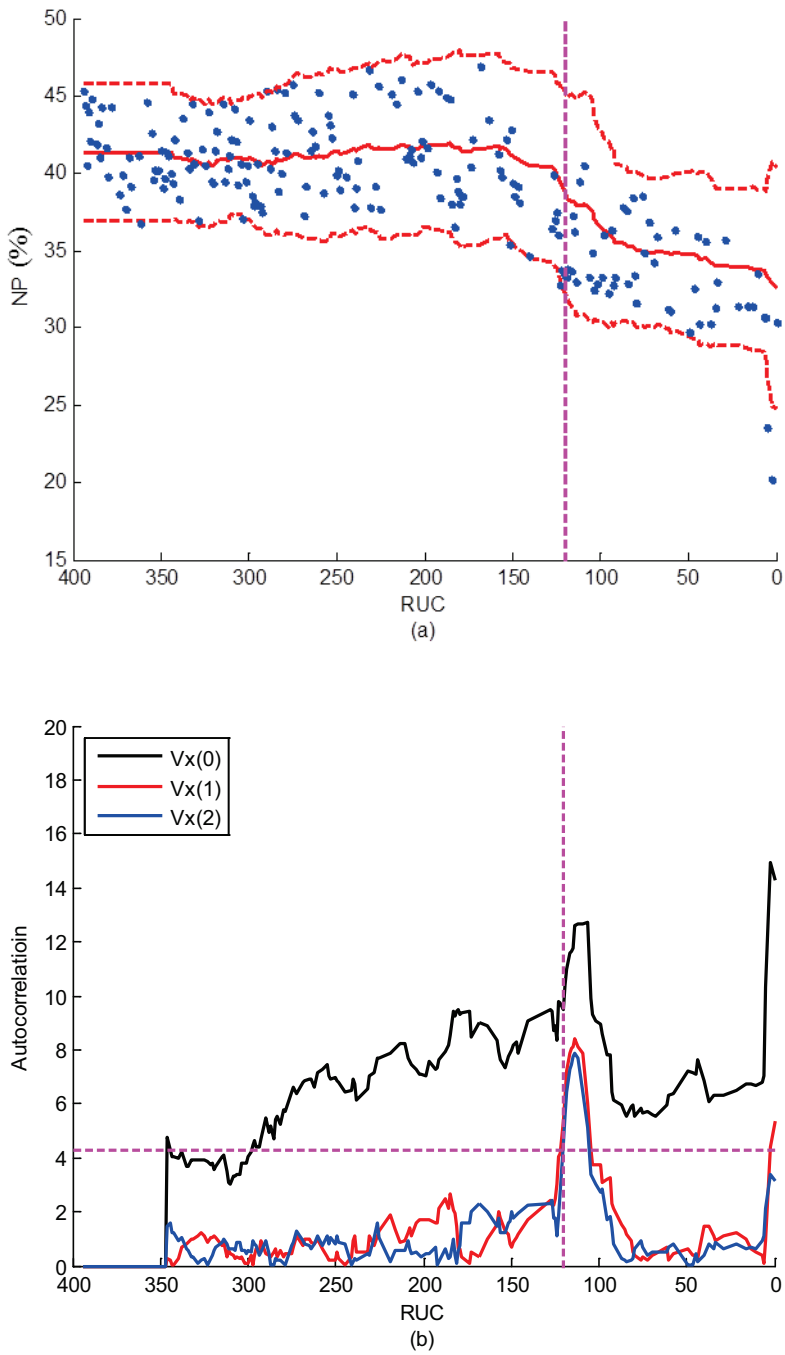


Figure 5.5 NP (a) samples, moving average, and moving standard deviation bounds (b) absolute values of moving experimental autocorrelations, the threshold, and the boundary line.

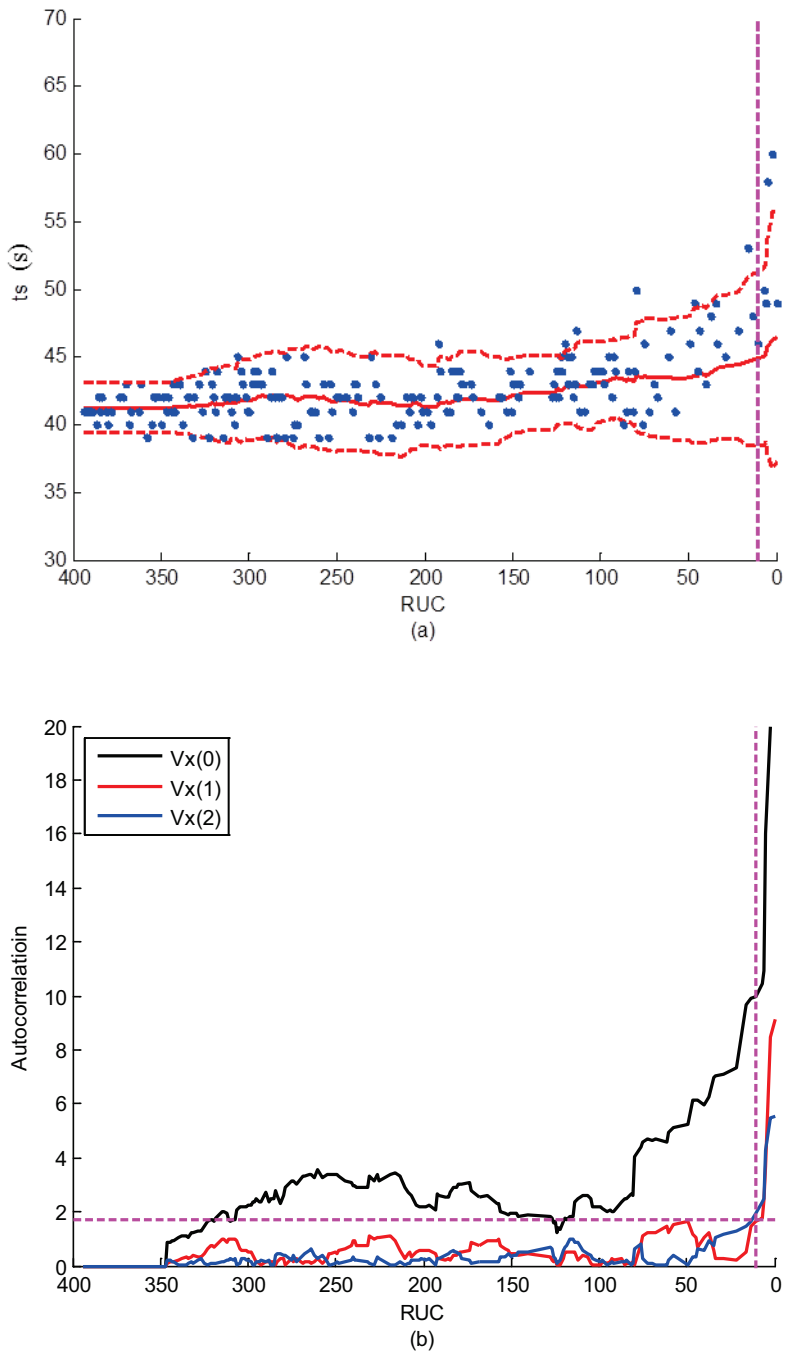


Figure 5.6 ts (a) samples, moving average, and moving standard deviation bounds (b) absolute values of moving experimental autocorrelations, the threshold, and the boundary line.

5.4 Offline evaluation

With the online classifier designed, the normal and degraded phases are then classified. For the same group shown in the previous section, the experimental autocorrelation $V_{EP}(\tau)$ and the power spectral density function $S_{EP}(j\omega)$ of the normal and degraded phases are plotted respectively in Figure 5.7.

As shown in the subplots (a) and (b), in the normal phase, $V_{EP_{nor}}(\tau)$ is a typical impulse function, while $S_{EP_{nor}}(j\omega)$ is almost an constant value over all frequency ranges. The results match with the expectations for a Gaussian white noise. As shown in the subplots (c) and (d), in the degraded phase, $V_{EP_{de}}(\tau)$ and $S_{EP_{de}}(j\omega)$ indicate that the underlying distribution is obviously different with that found in the normal phase.

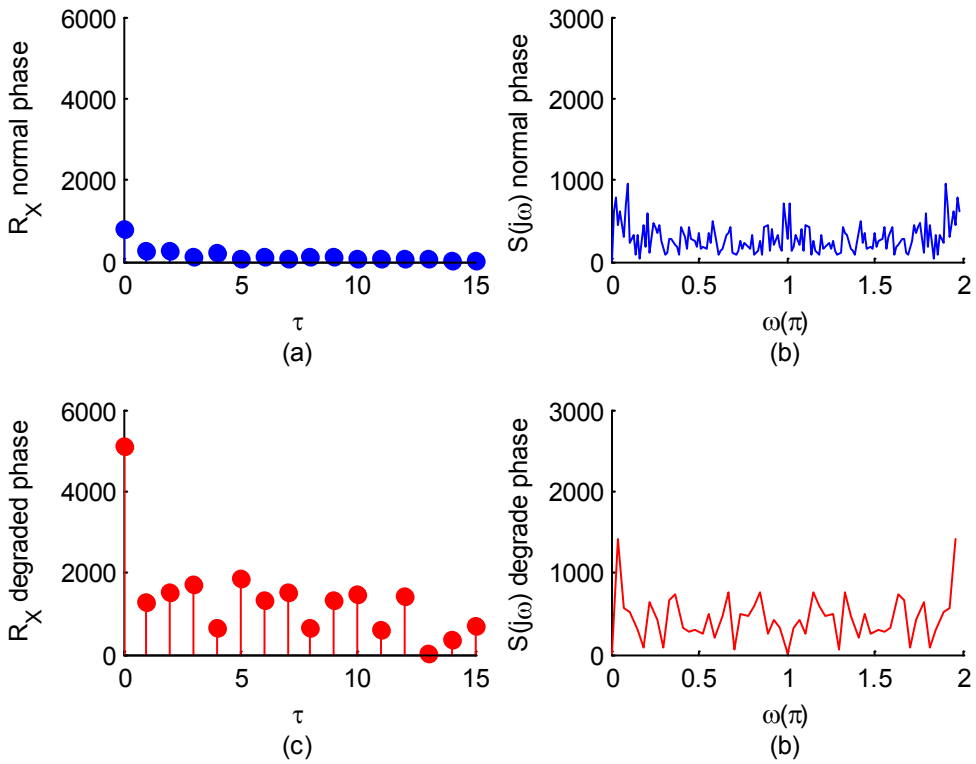


Figure 5.7 off line evaluation (a) $V_{EP_{nor}}(\tau)$ in normal phase, (b) $S_{EP_{nor}}(j\omega)$ in normal phase, (c) $V_{EP_{de}}(\tau)$ in degraded phase, (d) $S_{EP_{de}}(j\omega)$ in degraded phase.

The results of the online classifier applied on all 13 data groups have been lists in Appendix A. Some of the key parameters are listed in Table 5.1. The μ_{nor} and σ_{nor} indicates the underlying distribution of the corresponding random processes in the normal phase. The RUC_{bound} indicates the remaining useful cycles from the boundary line to the final failure "APU inability to start". It can be seen that these parameters varies significantly from one group to another.

Table 5.1 Parameters and boundary lines for all groups.

Group Number	$EP(C^o)$		$NP(\%)$		Threshold $V_{X,\tau=1,2}$	RUC_{bound}
	μ_{nor}	σ_{nor}	μ_{nor}	σ_{nor}		
1	543.78	22.26	41.065	2.638	362.00	124
2	532.92	18.86	41.143	2.6132	259.77	106
3	541.12	37.36	40.37	4.1336	1019.32	148
4	607.98	44.70	37.766	3.8024	1459.23	62
5	514.32	38.62	40.885	5.3483	1089.09	138
6	554.21	23.41	40.07	2.916	400.09	113
7	551.43	23.51	40.064	2.0238	403.80	54
8	535.69	32.02	39.312	4.6808	748.62	132
9	648.34	28.03	37.603	2.6879	573.74	115
10	511.94	20.32	42.008	2.9267	301.62	217
11	538.93	23.15	41.619	2.984	391.39	80
12	570.90	36.99	39.797	2.9943	995.37	77
13	582.98	48.57	37.072	3.9804	1722.96	39

The moving averages and data samples of EP and NP from all 13 groups are plotted together in Figure 5.8 and Figure 5.9 respectively. In order to make different groups comparable, each group is translated horizontally to let the boundary line overlap with y axis, and vertically to let the μ_{nor} overlap with x axis, that is a

translation from the coordinate (RUC, EP) to the coordinate $(RUC + boundary, EP - \mu_{nor})$. In the subplots (a) in the two figures, the moving averages illustrate that the starter degradations trends vary greatly in different groups. The subplots (b) in the two figures show that the online classifier has been effectively applied on all 13 groups.

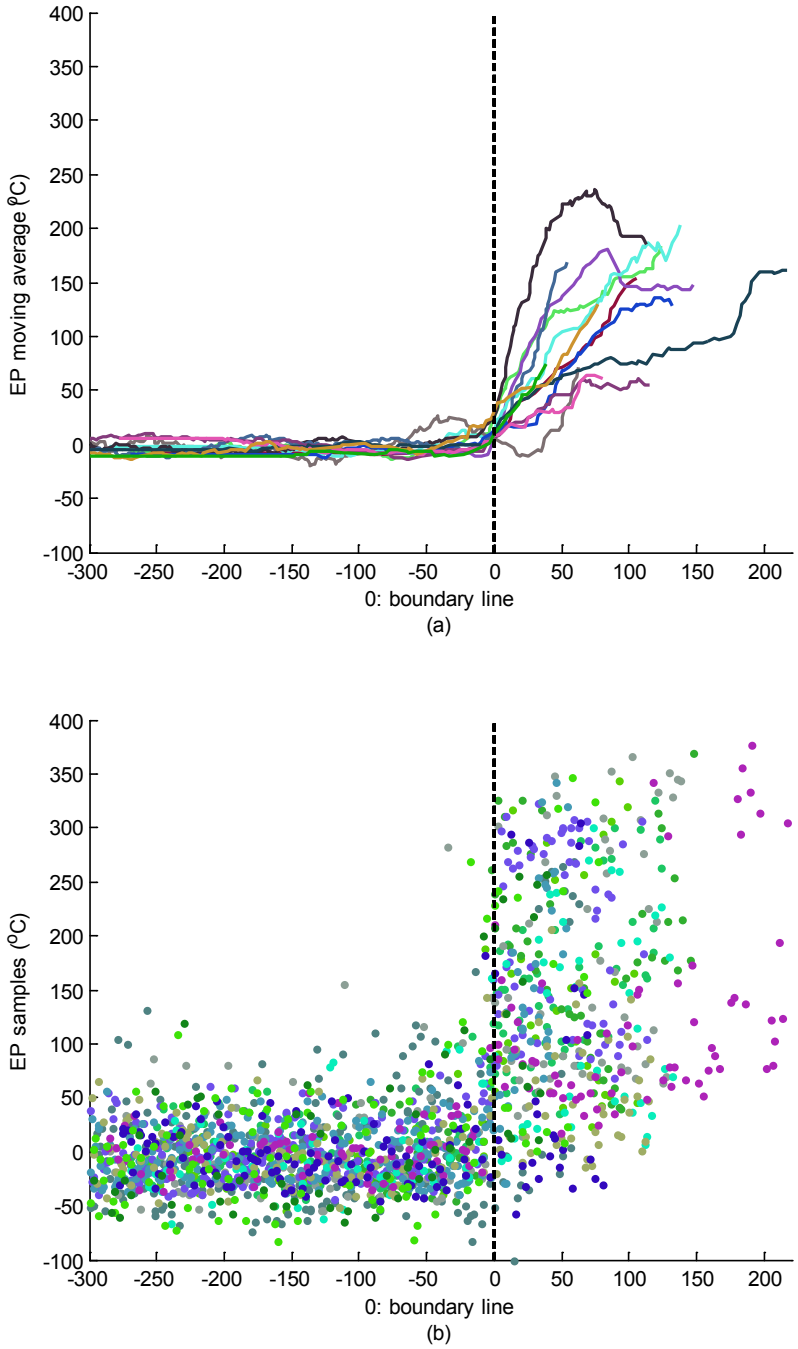


Figure 5.8 EP for all groups (a) moving average line, (b) samples.

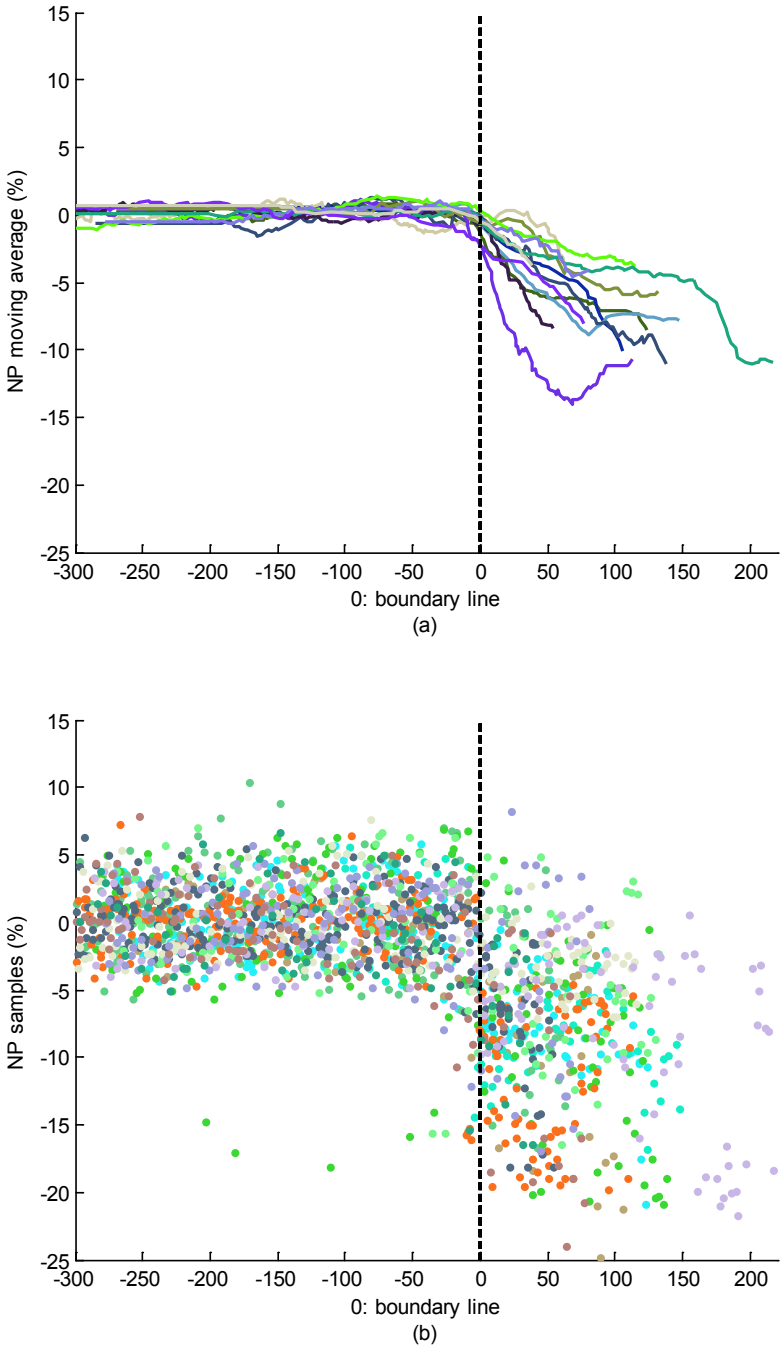


Figure 5.9 NP for all groups (a) moving average line, (b) samples.

5.5 Summary

In this chapter the condition monitoring data have been interpreted into random

processes based on physical understanding of gas turbine theory. Each independent group has been found to have two phases: the normal phase and the degraded phase. An online classifier has been developed to detect when the APU starter enters into the degraded phase.

When the designed classifier is applied to EP and NP respectively, it is found that the degraded phase of all 13 groups can be effectively detected. However, when the designed classifier is applied to ts, it is found that the results are not responsive effectively to reveal starter degradation. Therefore, the condition monitoring data is further simplified to the two variables EP and NP.

Based on the offline evaluation in Section 5.4, it is found that either the APU operating parameter in the normal phase or the degradation trend in degraded phase varies significantly from one group to another. Therefore, a general function of starter degradation that can be fit to all groups does not exist. In next chapter, a particle filtering approach will be implemented so that the prediction model can continuously update its parameters based on the measurements from each group in order to achieve accurate fault diagnostics and failure prognostics.

Chapter 6 Diagnostics and prognostics with particle filtering

Prognostics techniques analyze condition monitoring data prior to system failures in order to predict the future health conditions of targeted equipment based on information about their past usages and current states. Prognostics are more cost-effective than diagnostics because they can help prevent costly failures and reduce expenditures on condition-based maintenance.

This chapter presents a particle filtering (PF) approach applied on APU starter to achieve both diagnostics and prognostics simultaneously. The PF approach can identify the current starter health condition by filtering out engine uncertainty noises. Meanwhile, it can estimate the future degradation rate based on online condition monitoring data. The methodologies of recursive Bayesian estimation and the PF is introduced in Section 6.1. The proposed system state model of starter health condition and the PF implementation are demonstrated in Section 6.2. The PF results are analyzed in section 6.3.

6.1 Methodology

6.1.1 Recursive Bayesian estimation

Consider a general *system state model* in the form of Markov Process that

$$x_{k+1} = f_k(x_k, \omega_k), \tag{6.1}$$

$$y_k = h_k(x_k, v_k), \tag{6.2}$$

where k is the time step index, x_k is the unobserved *system state*, y_k is the observed time series *measurement*, and ω_k and v_k are the process noises of x_k and y_k respectively. The $f_k(\cdot)$ and $h_k(\cdot)$ are Markov process functions of system state and measurement.

The goal of a recursive Bayesian estimator is to make dynamic estimations of the system states x_k by constructing its conditional probability density functions (pdf) based on all the available measurements, $Y_k \triangleq [y_1, y_2, \dots, y_k]$. This conditional pdf is referred as *a posteriori* pdf, $p(x_k|Y_k)$. By contrast, the conditional pdf of x_k based on all measurements from the previous time step is referred as *a priori* pdf, $p(x_k|Y_{k-1})$ [67].

Suppose the a posteriori pdf at time step $k-1$ is available, that is $p(x_{k-1}|Y_{k-1})$, the a priori pdf at time step k can be estimated via the transition density function $p(x_k|x_{k-1})$ such that

$$p(x_k|Y_{k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|Y_{k-1})dx_{k-1}. \tag{6.3}$$

Then the a posteriori pdf at time step k is computed via the Bayes' theorem,

$$p(x_k|Y_k) = \frac{p(y_k|x_k)p(x_k|Y_{k-1})}{p(y_k|Y_{k-1})}, \tag{6.4}$$

where the normalizing constant is determined by

$$p(y_k|Y_{k-1}) = \int p(y_k|x_k)p(x_k|Y_{k-1})dx_k. \tag{6.5}$$

If the pdf of the initial state $p(x_0)$ is available, the dynamic estimations of the system state x_k can then be computed by iterations of the recursive process between the a priori pdf and the a posteriori pdf as shown in Figure 6.1.

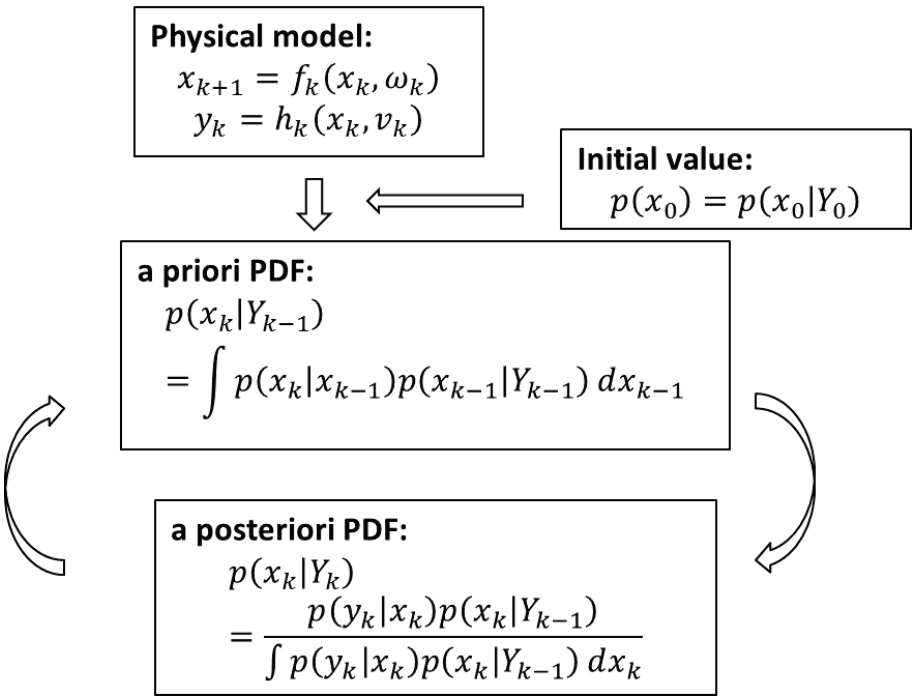


Figure 6.1 The recursive Bayesian estimator.

If the system is linear with independent Gaussian noises, the Bayesian state

estimator then reduces to the Kalman filter. If the system is nonlinear or has non-Gaussian noises, the analytical solutions are usually not available. In the past few decades, the extended Kalman filter (EKF) has become the most popular way for providing the approximate solutions. However, in the case of a system with severe nonlinearities, EKF will behave significant deviations from the true distribution. This deviation occurs because EKF relies on linearization to propagate the mean and covariance of the state [67].

6.1.2 Particle filtering

Sequential importance sampling (SIS) is the most basic Particle filtering (PF) algorithm for numerically implementing the recursive Bayesian estimator via Sequential Monte Carlo (SMC) simulations. The key idea of SIS is to represent the required a posteriori pdf $p(x_k|Y_k)$ by a set of random particles (samples) $x_k^i (i = 1,2,\dots,N)$ and their associated weights $w_k^i (i = 1,2,\dots,N)$.

$$p(x_k|Y_k) \approx \sum_{i=1}^{N} w_k^i \delta(x_k - x_k^i), \qquad \sum_{i=1}^{N} w_k^i = 1. \tag{6.6}$$

The w_k^i, normally known as *importance weight*, is the approximation of the probability density of the corresponding particle. In a nonlinear/non-Gaussian system where the state's distribution cannot be analytically described, the w_k^i of a dynamic

set of particles can be recursively updated through

$$w_k^i \propto w_{k-1}^i \frac{p(y_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, y_k)}, \tag{6.7}$$

where $q(x_k^i|x_{k-1}^i, y_k)$ is a proposal function called *importance density function*. There are various ways of estimating the importance density function. One common way is to select $q(x_k^i|x_{k-1}^i, y_k) = p(x_k^i|x_{k-1}^i)$ so that

$$w_k^i \propto w_{k-1}^i p(y_k|x_k^i). \tag{6.8}$$

In this particular case, the algorithm is also known as the bootstrap filter [68].

A common problem with SIS is known as the "degeneracy phenomenon". This occurs when the weight of one particle becomes significantly large while the weights of other particles degenerate to negligibility after a few iterations. One straight-forward way to reduce degeneracy effects is to perform the resampling procedure on the weights at every cycle so that they turn into uniformly distributed weights. Suppose the original N particles at the k time step iteration are termed as x_k^{i-}, for $i = 1,2, \dots N$. The resampling procedure would then be performed by the following two steps [69].

1. Generate a random number r that is uniformly distributed on $[0\ 1]$.
2. Accumulate the importance weights w_i into a sum, one at a time, until

$\sum_{i=1}^{j-1} w_i < r$ but $\sum_{i=1}^{j} w_i \geq r$. The new particle x_k^{i+} is then set equal to the jth original particle x_k^{j-}, that $x_k^{i+} = x_k^{j-}$ with probability w_j.

This resampling idea that the distribution presented by the set of new particles x_k^{i+} approaches to the a posterior pdf $p(x_k|y_k)$ as the number of particles N becomes sufficiently large is formally justified in [70].

6.2 Implementation

An accurate system state model is required before PF implementation. A brief review of the statistical properties of the condition monitoring data used to indicate the starter health condition (which were presented in Chapter 5) will be presented below for the purpose of establishing this system model.

The condition monitoring data consists of two random processes, EP and NP. They have similar statistical properties and high correlations with each other. Taking EP for example, the EP measurement is considered as the sum of two signals, the starter signal that indicates starter health condition and the noise signal caused by engine operating uncertainties.

Furthermore, an independent group is identified as having two phases, normal phase and degraded phase. In the normal phase, EP measurements satisfy a stationary Gaussian $\mathcal{N}(\mu_{nor}, \sigma_{nor}^2)$. The starter is healthy in this phase, and this healthy state is indicated by the starter signal which is a relative constant value equivalent to μ_{nor}. Meanwhile, the noise signal is a stationary white noise with

variance of σ_{nor}^2. In the degraded phase, EP measurements satisfy a non-stationary distribution that cannot be analytically described. The starter is experiencing degradation in this phase, and the degradation level is indicated by the starter signal which is the estimation of the measurements. Meanwhile, the noise signal is a non-stationary process with a variance that varies with the starter's degradation level.

Based on the above discussion, a PF approach supported by the proposed system state model should be able to achieve two goals. Firstly, the PF approach must filter out the white noise and identify the starter signal. Secondly, the PF approach must identify the degradation trend based on available measurements. This system state model is proposed as follows:

$$\overline{EP}_k: \quad x_{1_k} = x_{1_{k-1}} \left(\frac{x_{3_k}}{x_{3_{k-1}}} \right) \exp\left[x_{2_k} (RUC_k - RUC_{k-1}) \right], \tag{6.9}$$

$$\lambda_k: \quad x_{2_k} = x_{2_{k-1}} + \omega_{2_k}, \tag{6.10}$$

$$C_k: \quad x_{3_k} = x_{3_{k-1}} + \omega_{3_k}, \tag{6.11}$$

$$EP_k: \quad y_k = x_{1_k} + v_k. \tag{6.12}$$

where the subscript k represents the kth time step and RUC_k represents the starting cycle in this kth time step. There are three system states, $\overline{EP}$, λ, C, and one measurement, EP, in this system state model. These states and measurement are also denoted as x_1, x_2, x_3 and y respectively.

The first system state, $\overline{EP}$, represents the starter signal. As described in Equation

(6.9), its value at time step k is determined from the system states at the previous time step. The second system state λ represents the starter degradation rate. It is located in the exponential part of Equation (6.9). Therefore, the starter degradation rate between two adjacent starting cycles is indicated by e^{λ}. The higher λ is, the faster the starter degrades along an exponential growth. When $\lambda = 0$, no degradation develops between two starting cycles. The third system state C represents a discrete change of the starter degradation between two adjacent starting cycles. As can be seen in Equation (6.9), (C_k/C_{k-1}) is located in the product part. When $(C_k/C_{k-1}) = 1$, the starter degradation has no discrete change but develops continually along a the exponential curve e^{λ}. By contrast, when $(C_k/C_{k-1}) \gg 1$, the starter degradation develops suddenly and has a step upwards between two adjacent starting cycles. Both λ and C self-update with the random variables ω_{λ} and ω_C as described in Equation (6.10) and (6.11). These random variables are independent Gaussian white noise processes. The measurement used to update the above three system states is EP. As described in Equation (6.12), each EP measurement is the sum of the starter signal $\overline{EP}$ and the noise signal v, where the v represents the interference of engine operating uncertainties.

During the PF iterations, the systems states are estimated in the framework of recursive Bayesian by constructing their conditional pdf based on the measurements. The diagnostics is implemented by $\overline{EP}$ estimation that directly indicates the starter health condition. The discrete fluctuation of starter health condition is adapted by C estimation. And the prognostics is implemented by λ estimation. Once the

measurements stops, both λ and C are fixed with their most recent values. Thus the future degradation trend is expressed as an exponential growth of e^{λ}. The developed PF algorithm is summarized in Table 6.1:

Table 6.1 Particle filtering algorithm.

1. Generate N particles from the initial condition:

For $i = 1{:}N$

$$\begin{cases} x_{1_0}^{i+} = \mu_{nor} \\ x_{2_0}^{i+} \sim \mathcal{N}\left(0, \left(\omega_{2_k}\right)^2\right) \\ x_{3_0}^{i+} \sim \mathcal{N}\left(\mu_{nor}, \left(\omega_{3_k}\right)^2\right) \end{cases}$$

End for

2. Iteration between a priori and a posterior steps:

For $k = 1$: size of RUC

(a) Obtain a priori particles

For $i = 1{:}N$

$$\begin{cases} x_{2_k}^{i-} = x_{2_{k-1}}^{i+} + \omega_{2_k} \\ x_{3_k}^{i-} = x_{3_{k-1}}^{i+} + \omega_{3_k} \\ x_{1_k}^{i-} = x_{1_{k-1}}^{i+} \left(\dfrac{x_{3_k}^{i-}}{x_{3_{k-1}}^{i+}}\right) \exp\left[x_{2_k}^{i-}(RUC_k - RUC_{k-1})\right] \end{cases}$$

(b) Compute importance weights w_i for each particle $x_{1_k}^{i-}$

$$w_i = P\left[(y_k)\middle|x_{1_k}^{i-}\right] \propto \frac{1}{(2\pi)^{\frac{1}{2}} v_k} \exp\left(-\frac{v_k^2}{2}\right)$$

End for

(c) Scale the relative likelihoods w_{reg_i}

$$w_{reg_i} = \frac{w_i}{\sum_{i=1}^{N} w_i}$$

(d) Resample for posteriori particles $x_{j_k}^{i+} (j = 1,2,3)$

For $i = 1{:}N$

Generate a random number of $r \sim U[0,1]$

Find: $\sum_{i=1}^{j-1} w_{reg_i} < r$ & $\sum_{i=1}^{j} w_{reg_i} \geq r$

Set $x_k^{i+} = x_k^{j-}$ with probability w_{reg_i}.

End for

(e) Compute states estimations based on a posterior pdf $P[x_{j_k}|y_k]$

$$\begin{cases} \hat{x}_{1_k} = E(x_{1_k}|y_k) \\ \hat{x}_{2_k} = E(x_{2_k}|y_k) \\ \hat{x}_{3_k} = E(x_{3_k}|y_k) \end{cases}$$

End for

6.3 Results

6.3.1 Diagnostics results

The PF results of the two condition monitoring variables, EP and NP, are demonstrated in Figure 6.2 and Figure 6.3 respectively. In these figures, the x-axis value represents the starting cycles in terms of RUC with a left horizontal translation, such that the cycle 0 indicates the boundary line between the normal and the degraded phases determined in Chapter 5. Since the system state model only involves the step span, $RUC_k - RUC_{k-1}$, rather than the numerical value of RUC_k, this horizontal translation of x-axis does not affect PF results. As was mentioned before,

EP and NP has strong statistical dependence. A degraded starter is indicated by an increased EP and a decreased NP. Still taking EP as an example, the results are discussed as follows.

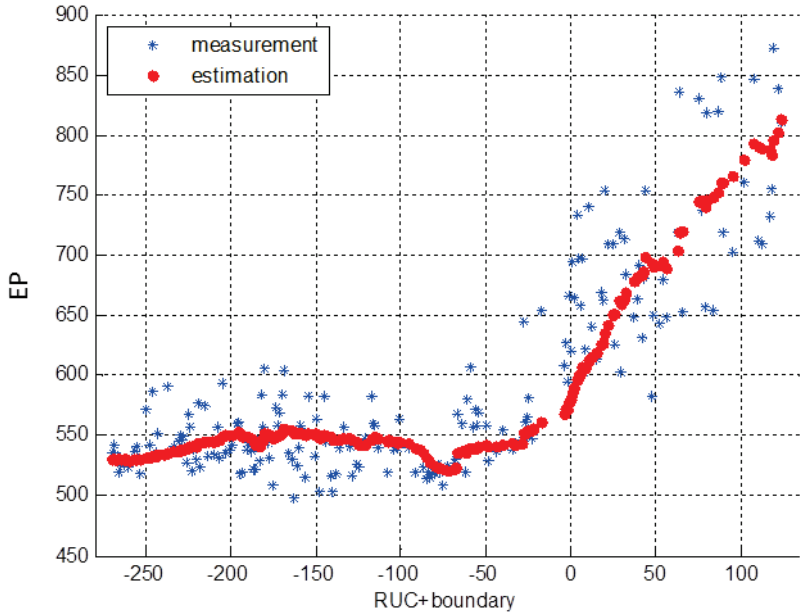
Figure 6.2 shows the estimations of the system states, $\overline{EP}$, λ, and C, in the full range of the group. In the subplot (a), $\overline{EP}$ is plotted as solid circle symbol that indicates the starter signal. EP is plotted as star symbols that indicates the original measurement. In the subplots (b) and (c), λ and C are plotted as solid circle symbol that indicate the starter condition's degradation rate and sudden fluctuation respectively. Comparing the subplots (a), (b) and (c), the starter degradation in this group appears to have experienced three phases. Each of the phases is described as follows.

The first phases occurs in the zone $x < -15$. Since $x = 0$ is the boundary line determined by the online classifier, this phase has been identified as normal phase. Also as shown in the figure, all the three system states maintain stable values. $\overline{EP} \approx \mu_{nor}$ indicates that the starter is in a good health condition. $\lambda \approx 0$ indicates that there is no degradation developed. And $C_k \approx C_{k-1}$ indicates that there is no sudden fluctuation of starter health condition between two adjacent starting cycles.

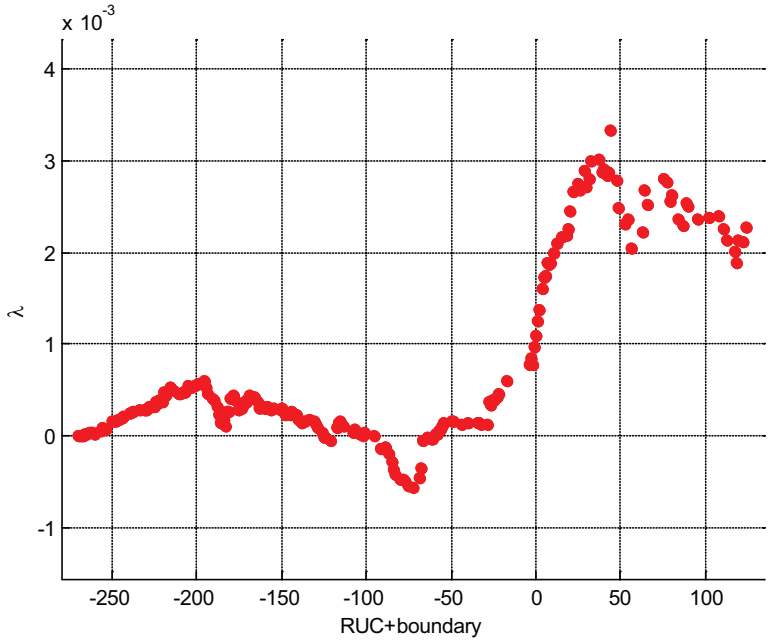
The second phase occurs in the zone $-15 < x < 30$. It is a transitional phase from normal to degradation. $\overline{EP}$ indicates that the starter degradation develops rapidly. λ indicates that the degradation rate grows from e^0 to around $e^{0.003}$. C indicates that the degradation does not follow a smooth exponential growth but jumps upwards suddenly from one starting cycle to the next.

The remaining phase is in the zone $x > 30$. The degradation becomes relatively stable. $\overline{EP}$ indicates that the starter continuous to degrade. However, λ indicates that the degradation rate stops increasing but fluctuates repetitively between $e^{0.002}$ to $e^{0.003}$. C indicates that the degradation's upwards jumping stops. But the starter health condition is more unstable than that in the first phase.

In conclusions, the diagnostics of the APU starter health condition has been implemented by the proposed PF approach. Firstly, the subplot (a) shows that the noise signal has been effectively filtered out. $\overline{EP}$ is a relatively smooth and continuous curve that can responsively indicate the starter health condition. Secondly, the subplots (b) and (c) show that λ and C have self-updated dynamically to indicate the progressive rate of starter degradation.


measurement
estimation
EP
900
850
800
750
700
650
600
550
500
450
-250
-200
-150
-100
-50
0
50
100
RUC+boundary

(a)


x 10^-3
λ
4
3
2
1
0
-1
-250
-200
-150
-100
-50
0
50
100
RUC+boundary

(b)

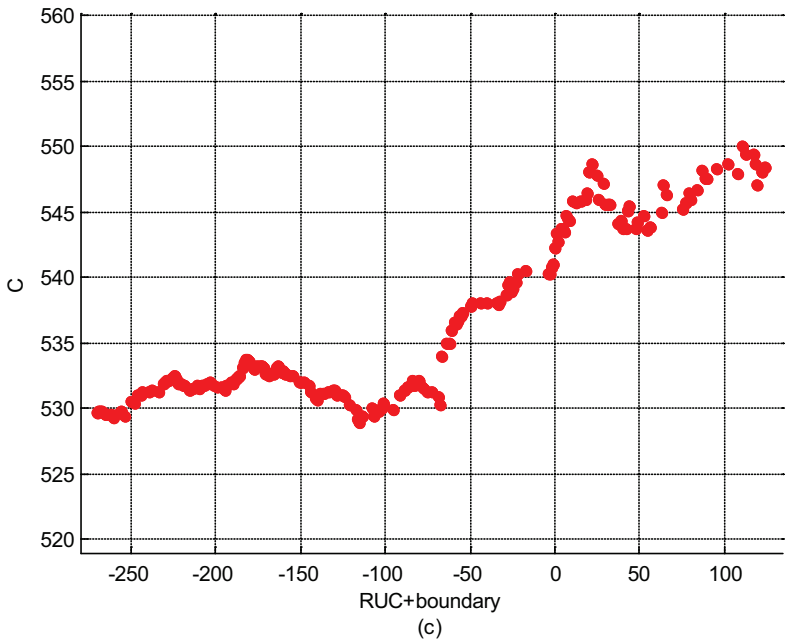
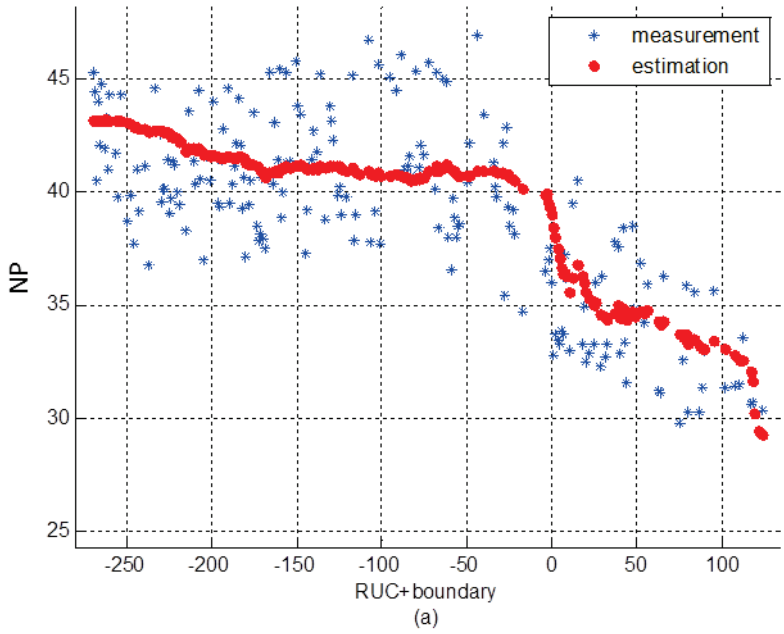
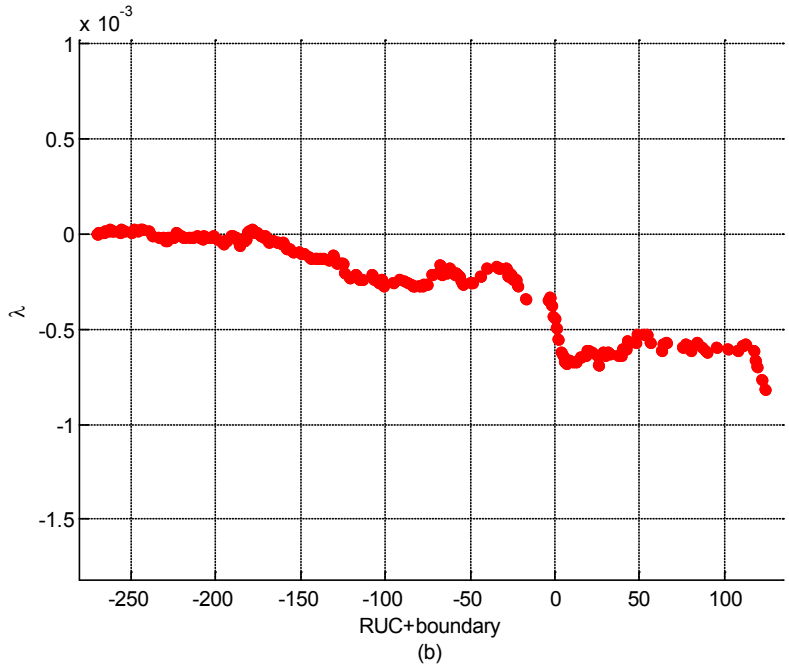


Figure 6.2 (a) $\overline{EP}$ estimation, (b) λ estimation, (c) C estimation.





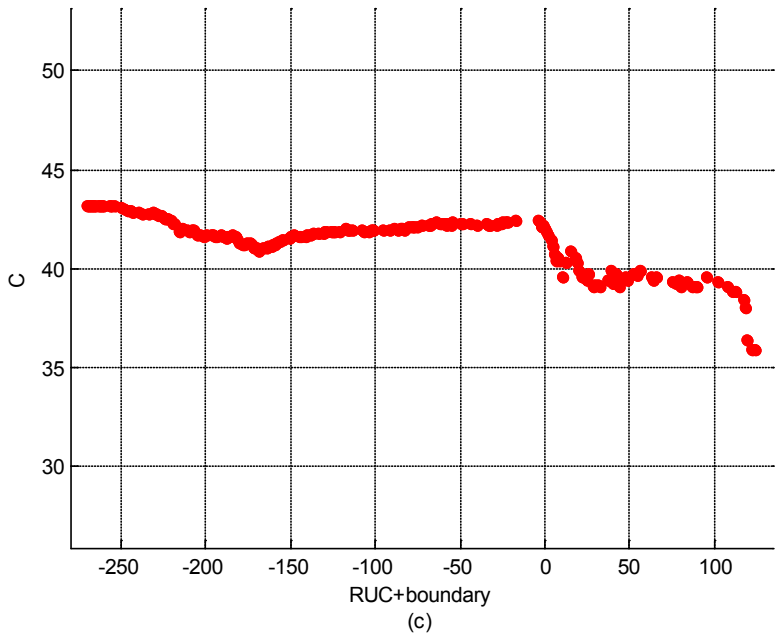


Figure 6.3 (a) $\overline{NP}$ estimation, (b) λ estimation, (c) C estimation.

Once the $\overline{EP}$ and $\overline{NP}$ values for a particular starting cycle have been determined, the corresponding measurement, EP and NP, can then be expressed as the pdf of a 2-D Gaussian distribution by considering them as a 2-dimensional vector

$x = (EP, NP)$, so that

$$f_x(EP, NP) = \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\},$$

$$\mu = (\overline{EP}, \overline{NP}),$$

$$\Sigma = \begin{pmatrix} v_{EP}^2 & \rho v_{EP} v_{NP} \\ \rho v_{EP} v_{NP} & v_{NP}^2 \end{pmatrix}, \tag{6.13}$$

where ρ has been calculated in Section 5.1 with $corr(EP, NP) = -0.8881$. Taking the starting cycle $x = 100$ in Figure 6.2 and Figure 6.3 for instance, the corresponding $\overline{EP} = 800$ and $\overline{NP} = 33$. In addition, the corresponding $v_{EP}^2 = 4937$ and $v_{NP}^2 = 6.5$ have been determined from Figure 5.4 and Figure 5.5. With all these known values, the joint pdf of EP and NP is then described in Figure 6.4.

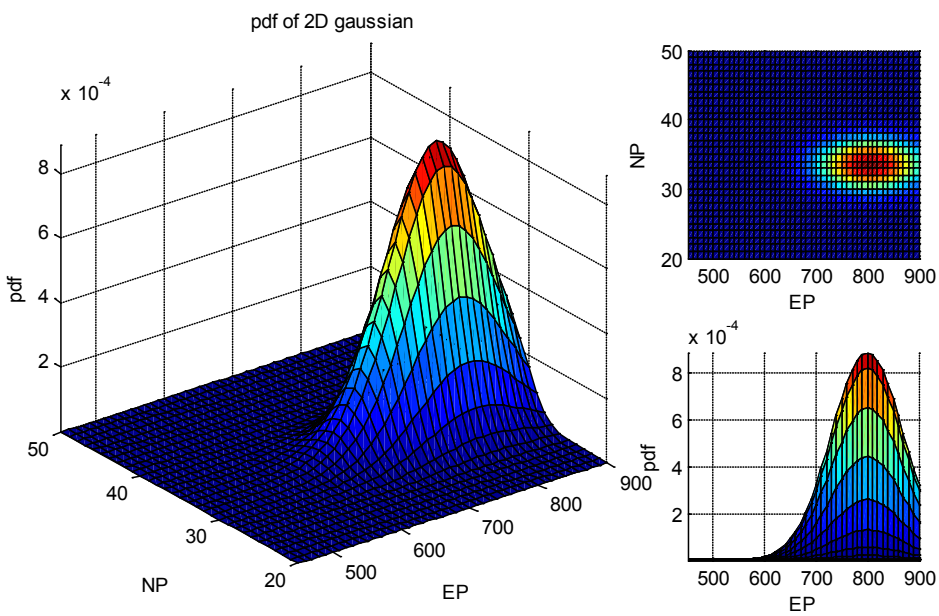


Figure 6.4 Joint distribution of $(EP\ NP)$

6.3.2 Prognostics results

The basic idea of prognostics on the APU starter is that λ are fixed at its most recent values updated by the available measurements. Then the future degradation trend is expressed as an exponential growth, e^{λ}, started from the latest $\overline{EP}$ or $\overline{NP}$ estimations. Figure 6.5 and Figure 6.6 show the PF results when the prognostics is triggered at 100 and 50 starting cycles prior to the reset event respectively. Furthermore, since there are totally 13 groups have been identified as containing starter degradation information, the processing results for the other 12 groups using the developed PF based approach are plotted in Appendix A, in which the prognostics are triggered at the last 30 cycles prior to the reset event.

As shown in these figures, this proposed PF approach has been validly implemented on APU starter prognostics. In a single group, as more measurements become available over the period of state estimation, the PF can self-update the system states and generate more accurate forecasting results. From different groups, as the corresponding measurement trends vary with each other, the PF can self-adapt to each trend and generate the most pertinent predictions.

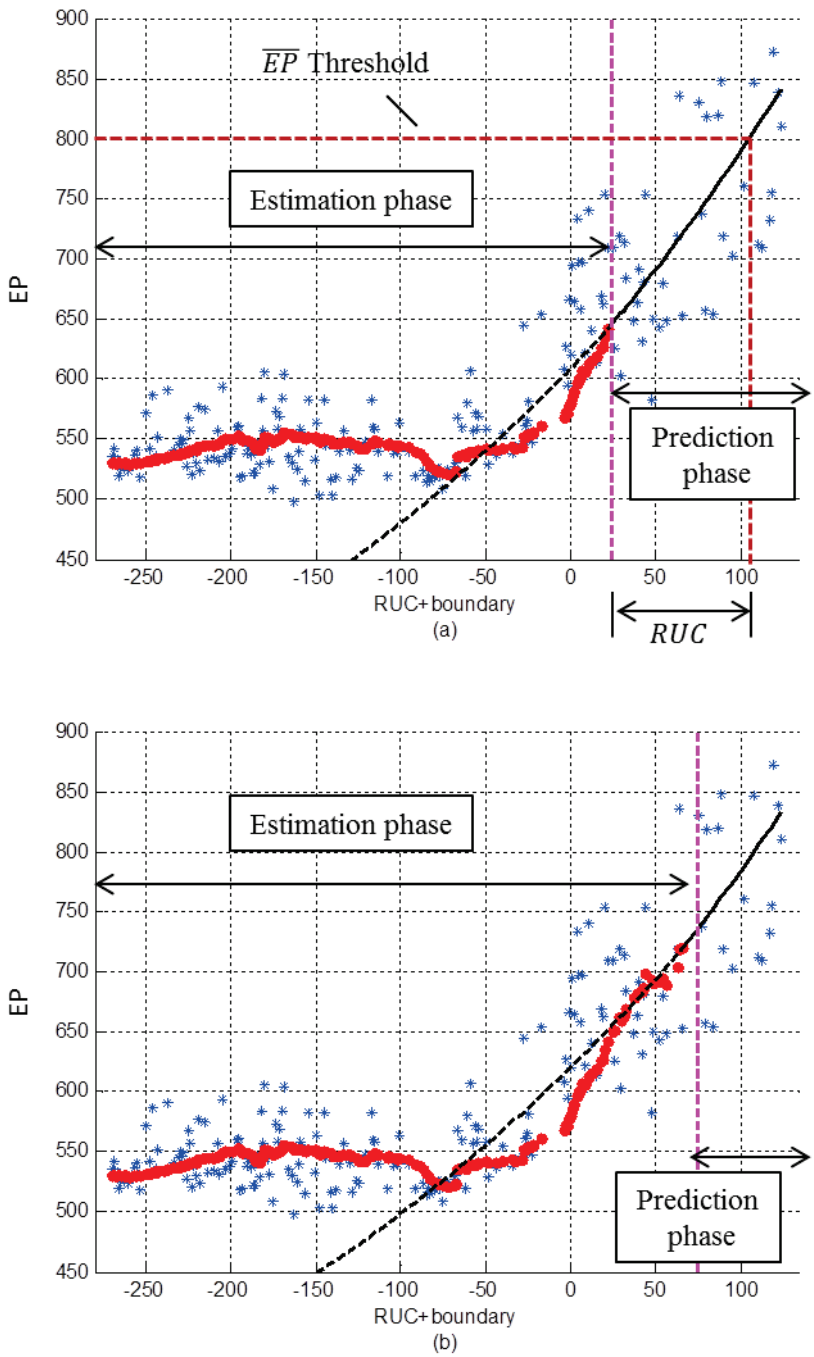


Figure 6.5 PF prognostics results of EP

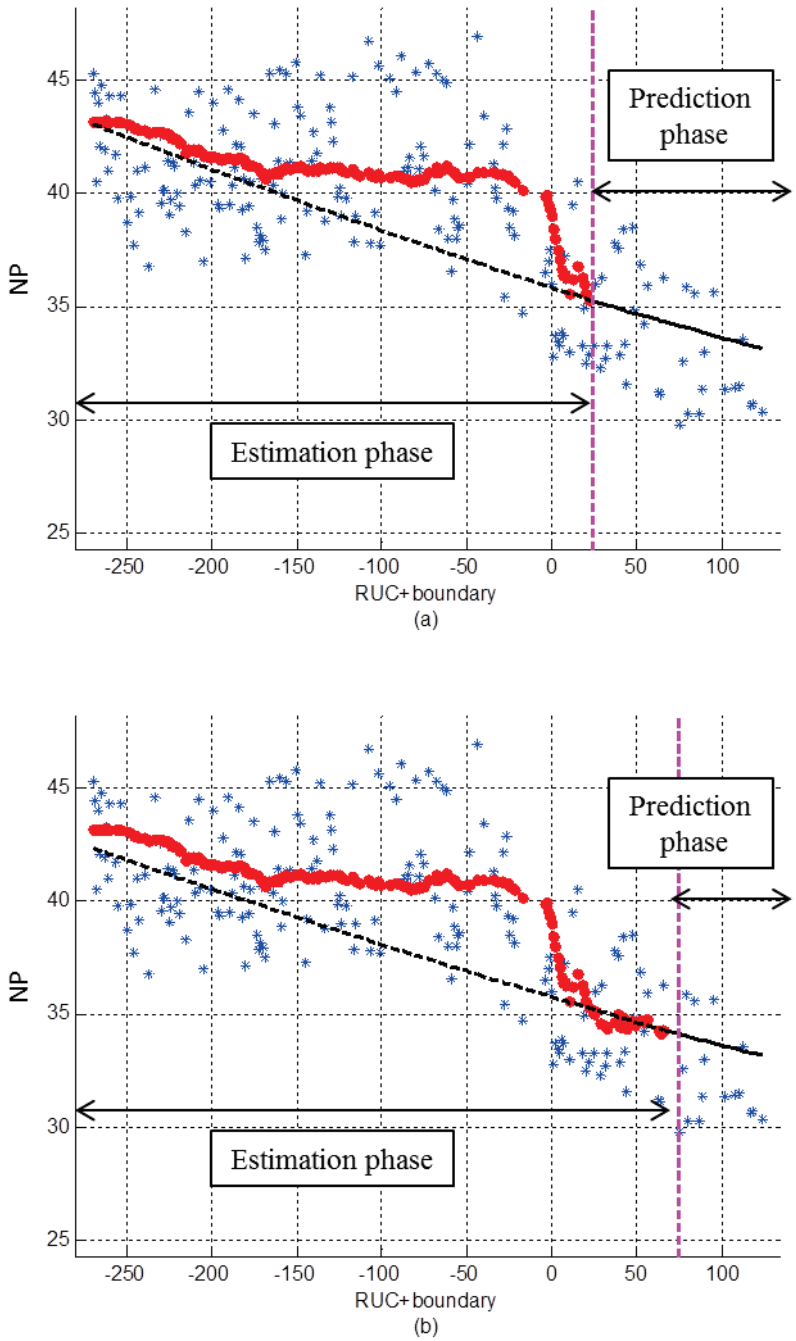


Figure 6.6 PF prognostics results of NP.

An important concept in Prognostics is the remaining useful life (RUL), which is also referred as RUC in this thesis. It indicates the time left before the final failure of target equipment. A common method to decide RUL is to select a threshold of the

feature that indicates the target equipment's health condition. Once the predicted value reaches the threshold, the corresponding time span is then determined as the RUL. Taking Figure 6.5 for example, suppose a threshold of $\overline{EP}$ is selected 800^oC, the prediction curve trigged at around $x = 25$ crosses the threshold at $x = 105$, the corresponding time span $105 - 25 = 80$ is then determined as the RUC.

However, because of the intrinsic properties of APU and its starter, both APU operating parameters and starter degradation conditions vary significantly from one group to another. Therefore, a $\overline{EP}$ or $\overline{NP}$ threshold that applies to all groups for RUC prediction can be hardly selected. This is demonstrated in the Appendix A, Table 6.2, and Figure 6.7.

Table 6.2 lists the key parameters from all groups. As can be seen, for different groups, μ_{nor} is in a range from 511^oC to 648^oC, indicating that the mean values of maximum exhaust gas temperatures in starting processes with healthy starter may vary more than 100^oC. Also, RUC_{bound} is in a range from 39 to 217, indicating that the time spans from initial degradation to finial failure may vary more than 180 cycles. Last, $\overline{EP}_{end} - \mu_{nor}$ in Table 6.2 is plotted in Figure 6.7, which is in a range from 61^oC to 289^oC, indicating that the starter degradation levels at the final failures do not have a general standard threshold.

The varieties discussed above is reasonable, because APU is a complex system that involves manufacturing, assembling, and operating of hundreds of components, also because the starter degradation involves complex conditions of aging, operating and environment. Therefore, although the PF can effectively estimate and predict the

starter health condition, a precise RUC for a particular APU starter is still difficult to be determined.

Table 6.2 PF results for all groups.

Group	μ_{nor}	RUC_{bound}	EP_{end}	$\overline{EP}_{end}$	$EP_{end} - \mu_{nor}$	$\overline{EP}_{end} - \mu_{nor}$
1	543.78	124	810.12	806.57	266.3479	262.79
2	532.92	106	588.01	721.47	55.0997	188.55
3	541.12	148	913.64	758.38	372.5201	217.26
4	607.98	62	838.58	766.65	230.6016	158.67
5	514.32	138	863.37	775.61	349.0587	261.29
6	554.21	113	662.17	769.13	107.9696	214.92
7	551.43	54	872.83	840.98	321.4063	289.55
8	535.69	132	612.02	670.61	76.3328	134.92
9	648.34	115	736.08	709.91	87.7410	61.57
10	511.94	217	817.61	742.44	305.6733	230.50
11	538.93	80	507.91	663.65	-31.0121	124.72
12	570.90	77	847.13	733.23	276.2319	162.33
13	582.98	39	845.12	654.27	262.1456	71.29

- μ_{nor}: Average value of EP in normal phase.
- RUC_{bond}: The RUC at the boundary line between normal and degraded phase.
- EP_{end}: EP at the last cycle when APU "inability to start" occurs.
- $\overline{EP}_{end}$: $\overline{EP}$ at the last cycle estimated by PF.

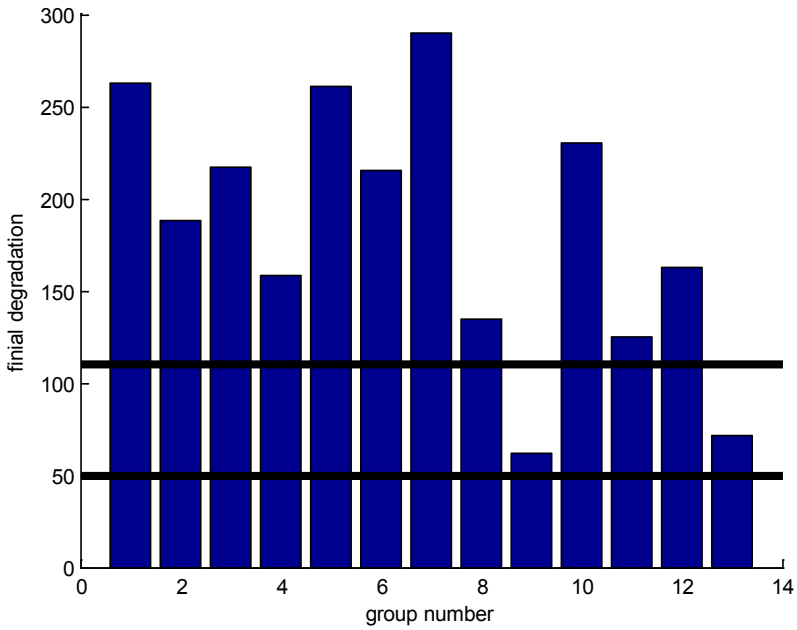


Figure 6.7 $(\overline{EP}_{end} - \mu_{nor})$ bars for all groups.

However, from the view point of actual airline operation, the aim is to prevent unnecessary economical loss and the potential flight risk. Examining Figure 6.7, it shows that $\overline{EP}_{end} - \mu_{nor} = 50^oC$ is a safe threshold for all groups, while 100^oC is a safe threshold for most of the groups. Therefore, this thesis suggest to perform a maintenance action once the estimated $\overline{EP}$ is 50^oC higher than μ_{nor}, or once the starter is determined as entering into the degraded phases by the classifier designed in Chapter 5. Although the APU starter may still support another dozens of starting cycles, EP and NP have already obviously deviated from normal standards. These deviations will impair the APU engine core, shorten the APU life, increase the starting surge and starting overheat risks and endanger the aircraft flight.

6.4 Summary

In this chapter, a PF based approach with the associated system state model has been proposed based on an understanding of the data's statistical properties and gas turbine theory. Three states, $\overline{EP}$ (or $\overline{NP}$), λ, and C, have been estimated in the Bayesian framework based on the available condition monitoring data, EP and NP, to indicate the current starter health condition and future degradation trends.

This proposed PF approach has been validly applied on the APU starter diagnostics and prognostics. Firstly, the noise signal caused by engine operating uncertainties has been effectively filtered out. The starter signal that indicates the starter health condition has been clearly described. Secondly, these system states can dynamically self-update based on measurements from different phase or different group and generate most effective estimations and predictions.

Because of the tremendous varieties exhibited in different groups, the determination of general thresholds of $\overline{EP}$ and $\overline{NP}$ for all the APU RUC prediction is very challenging. However, from the actual airline operation viewpoint, this thesis suggests that a maintenance action once the estimated $\overline{EP}$ is 50^oC higher than μ_{nor}, or once the starter is determined as entering into the degraded phases by the designed online classifier. This criterion can effectively support the CBM decision-making to prevent APU damage and enhance flight safety.

Chapter 7 Conclusions and future works

7.1 Conclusions

APU starter is a crucial component that supports APU operation and aircraft operation. Starter performance degradation can cause significant economic loss and raise aircraft flight risk. However, the condition-based maintenance (CBM) for the APU starter has not been well investigated. Accordingly, this thesis proposed a framework to support the entire CBM procedure, from data acquisition to data processing to decision-making.

Firstly, the dataset collected from actual airline operations has been corrected to the ISA sea level ambient conditions following the principle of Mach number similarity. After that, two essential parameters, the corrected EGT peak value during starting process, EP, and the associated rotational speed, NP, are selected as the condition monitoring data. The physical relationship between EP, NP and the starter assistance power, $P_{starter}$, has been derived from general gas turbine engine theory. Based on this relationship, 13 groups of the dataset have been identified as informative data source for estimating the starter degradation. Next, EP and NP have been treated as two random processes in terms of RUC. The techniques, kernel smoother and moving average, have been applied for the statistical properties analyses. With the statistical properties, an online classifier based on the moving

autocorrelations has been designed that can effectively detect the starter's initial degradation. Finally, a PF based approach with associated system state model has been developed and implemented for the APU starter fault diagnostics and failure prognostics.

Throughout the entire thesis, a complete CBM for APU starter in civil aviation field has been implemented. There are three key methodologies to support this CBM framework:

1. The EP and NP indicator for APU starter health condition.
2. The autocorrelation classifier for APU starter degradation detection.
3. The PF estimator for APU starter fault diagnostics and failure prognostics.

7.2 Future works

There are still many challenges that need to be addressed in the future research works. These works can be conducted on the following subjects:

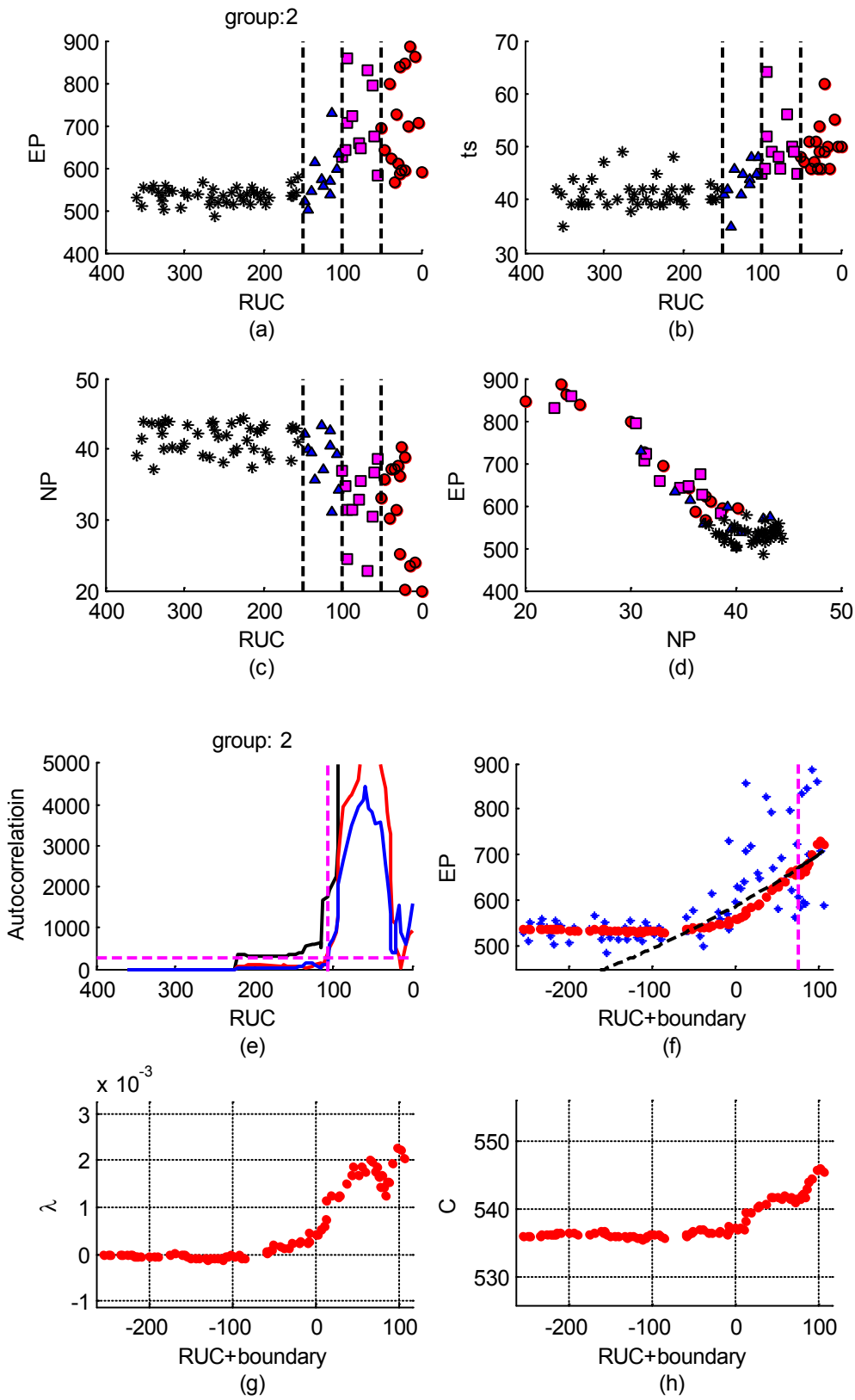
- The proposed CBM framework will be applied to other mechanical systems such as aircrafts' main engines.
- The dataset will be further explored to extract other relevant features for starter health condition indication.
- In the case that a particular gas turbine engine's performance characteristics (performance maps) at design level are accessible, the engine operating uncertainties can be determined more precisely.

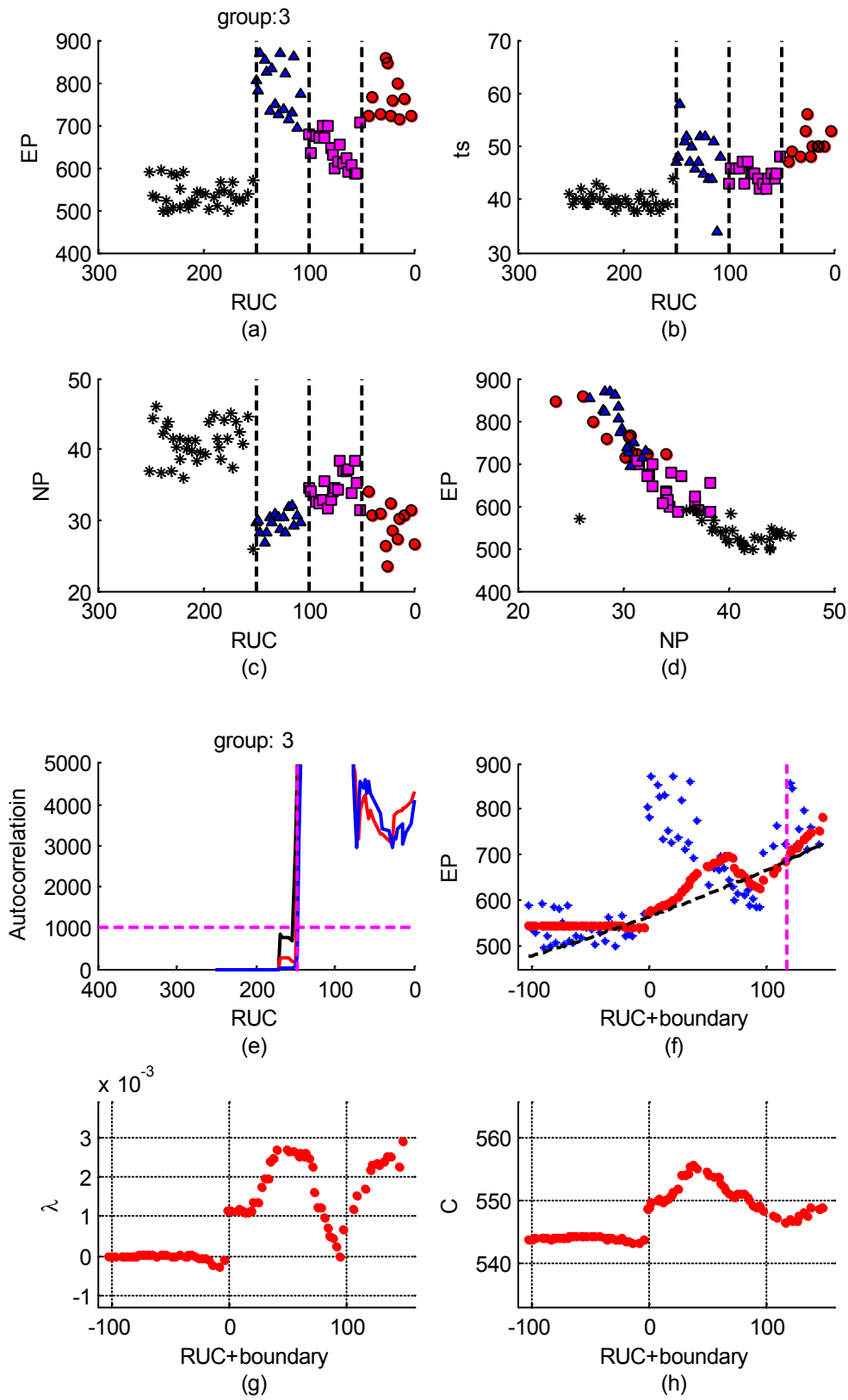
- In this thesis, the identification of different failure modes caused by different components has been made based on offline data observation. The future research aims to implement an online automatic identification.
- In this thesis, the prognostics assume that starter degradation follows a certain exponential growth pattern. It cannot effectively handle the repetitive fluctuation of the starter degradation. The future research should consider to combine the data-driven techniques such as neural fuzzy and the PF applied in this thesis. This combination may achieve better results in failure prognostics.

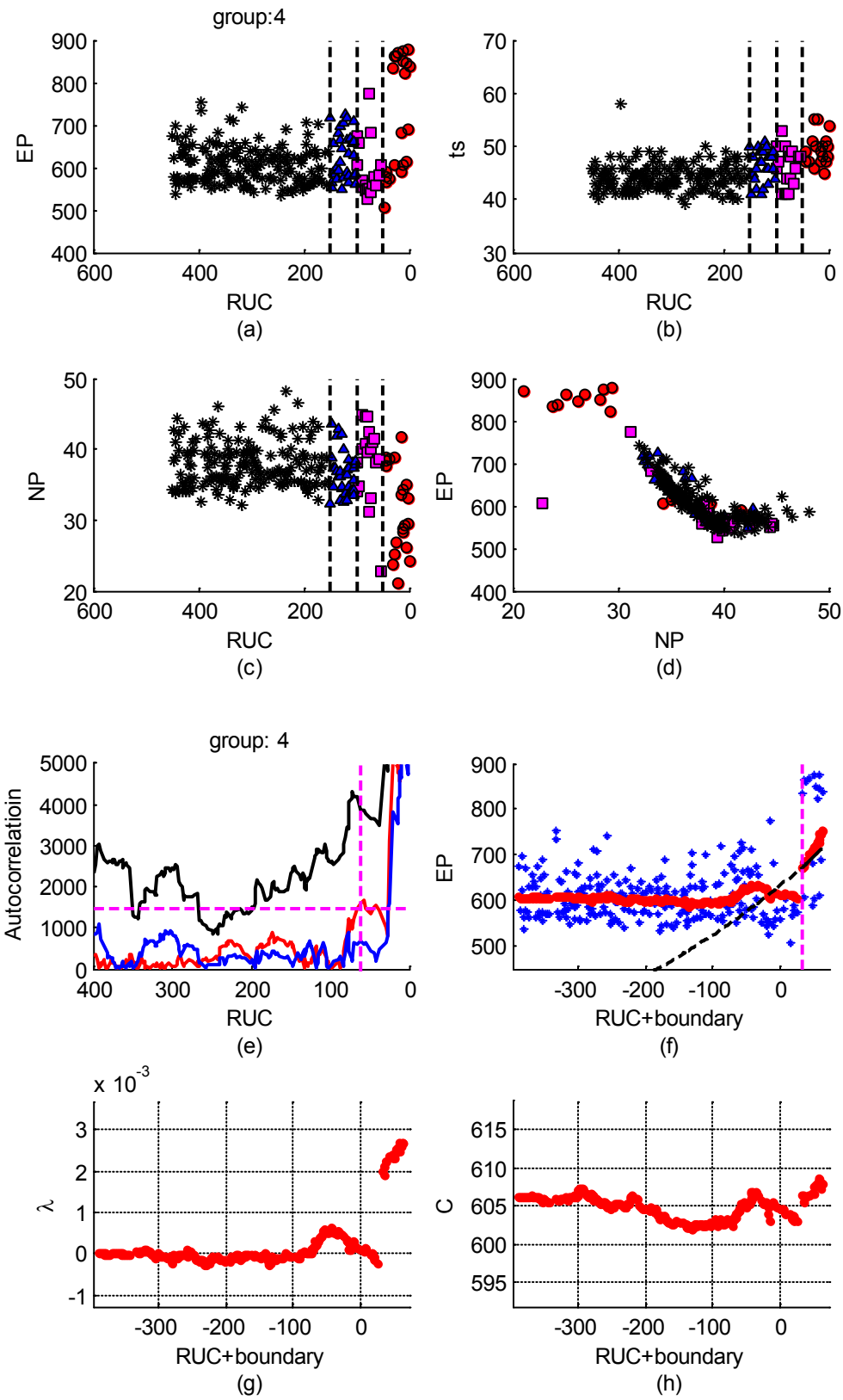
Appendix A Results for all starter degradation groups

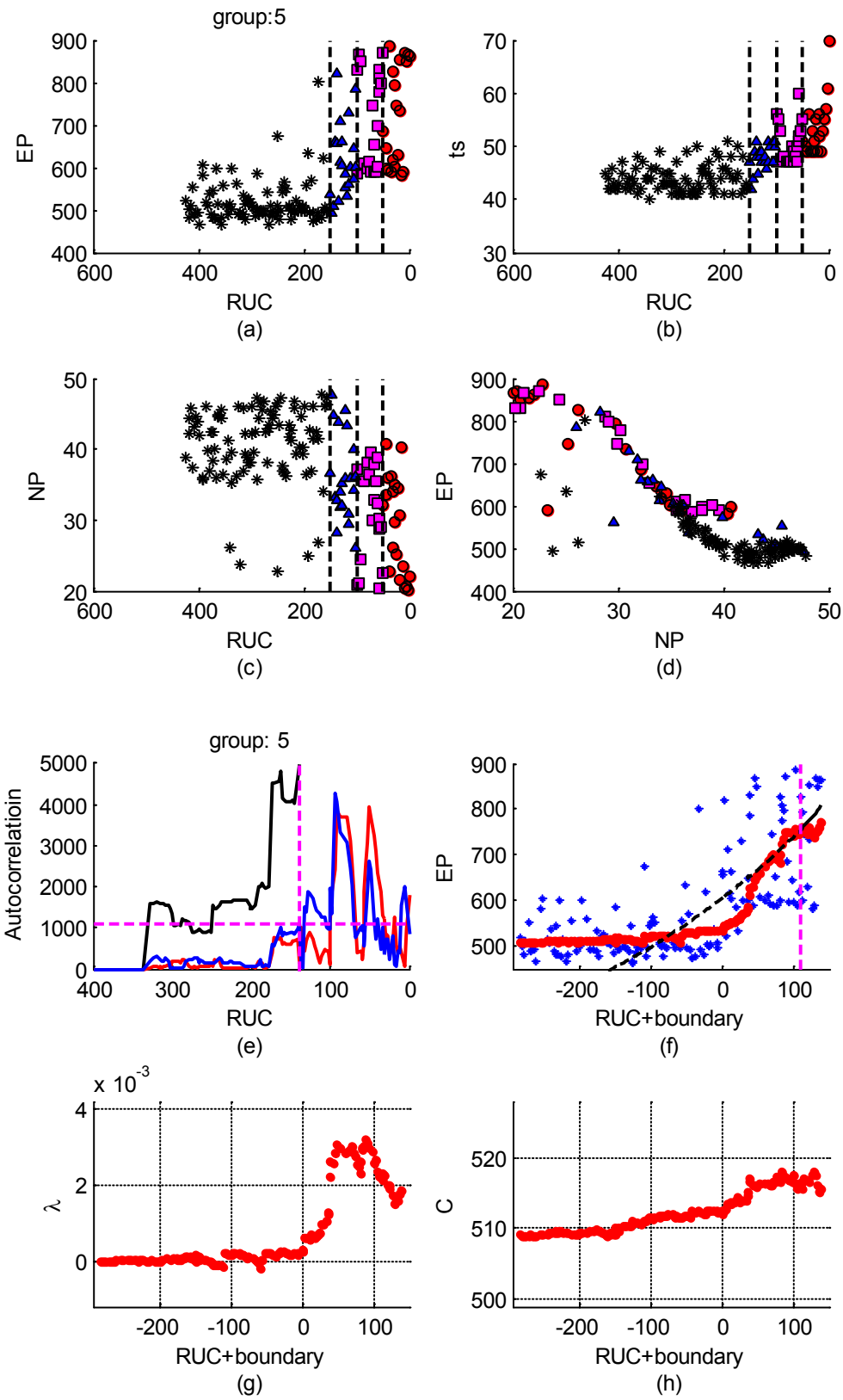
Appendix A demonstrates data patterns, online classifier and PF estimations of the groups relevant to APU starter degradation. Each group is described by the following figures:

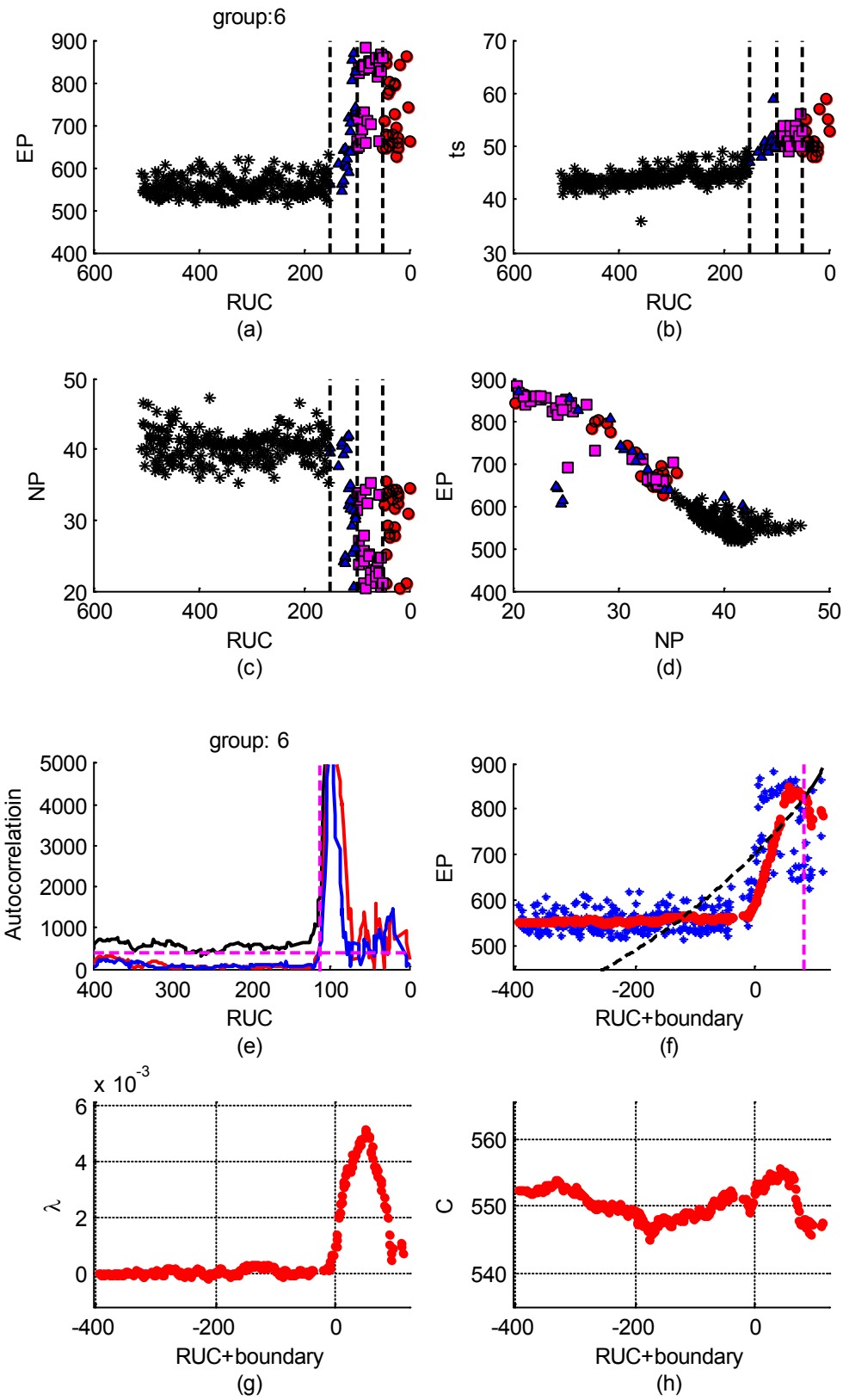
- Subplot (a): EP against RUC.
- Subplot (b): ts against RUC.
- Subplot (c): NP against RUC.
- Subplot (d): EP against NP.
- Subplot (e): online classifier based on autocorrelation.
- Subplot (f): $\overline{EP}$ diagnostics and $\overline{EP}$ prognostics at last 30 cycles.
- Subplot (g): λ diagnostics.
- Subplot (h): C diagnostics.

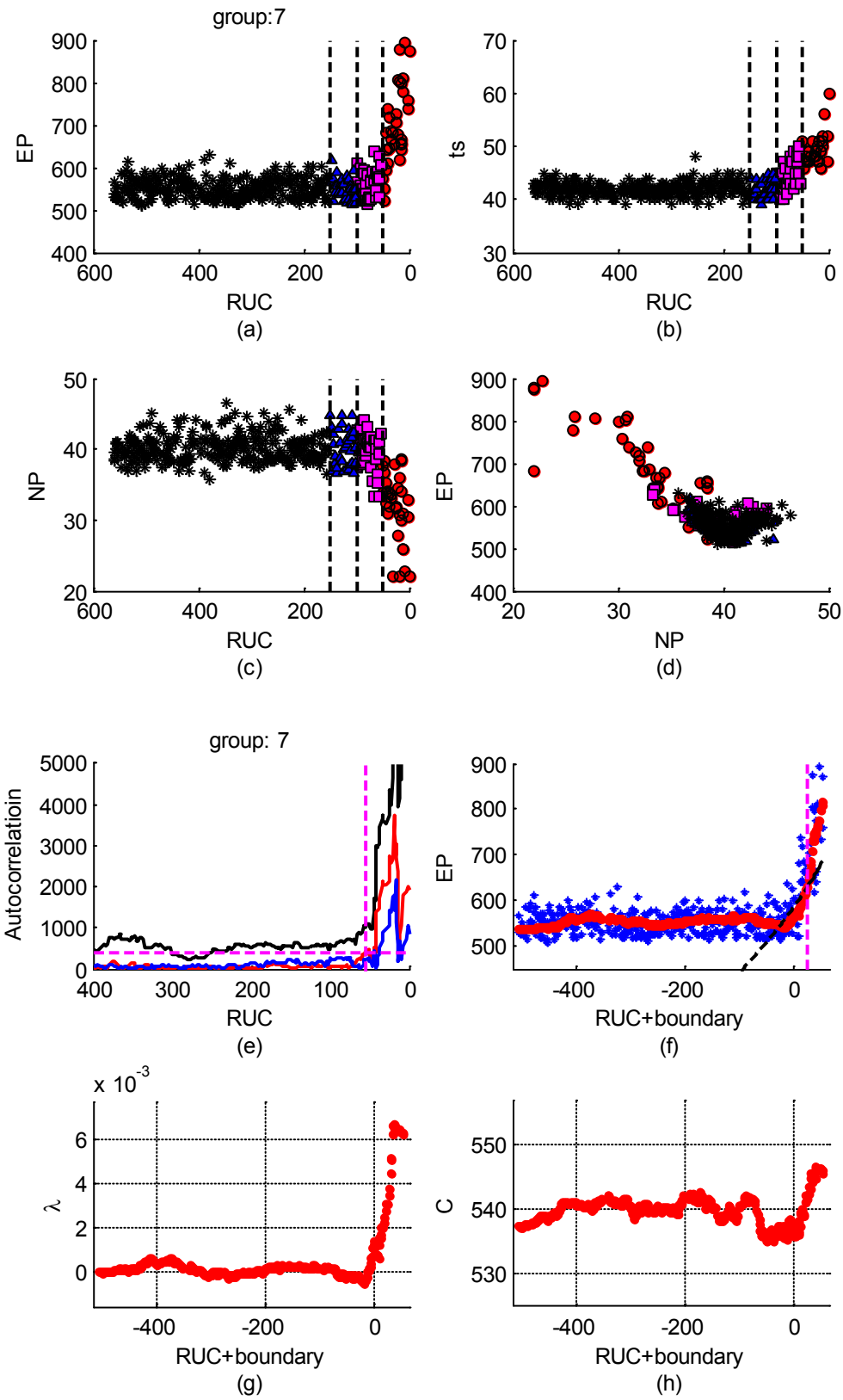

group:2
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 2
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)

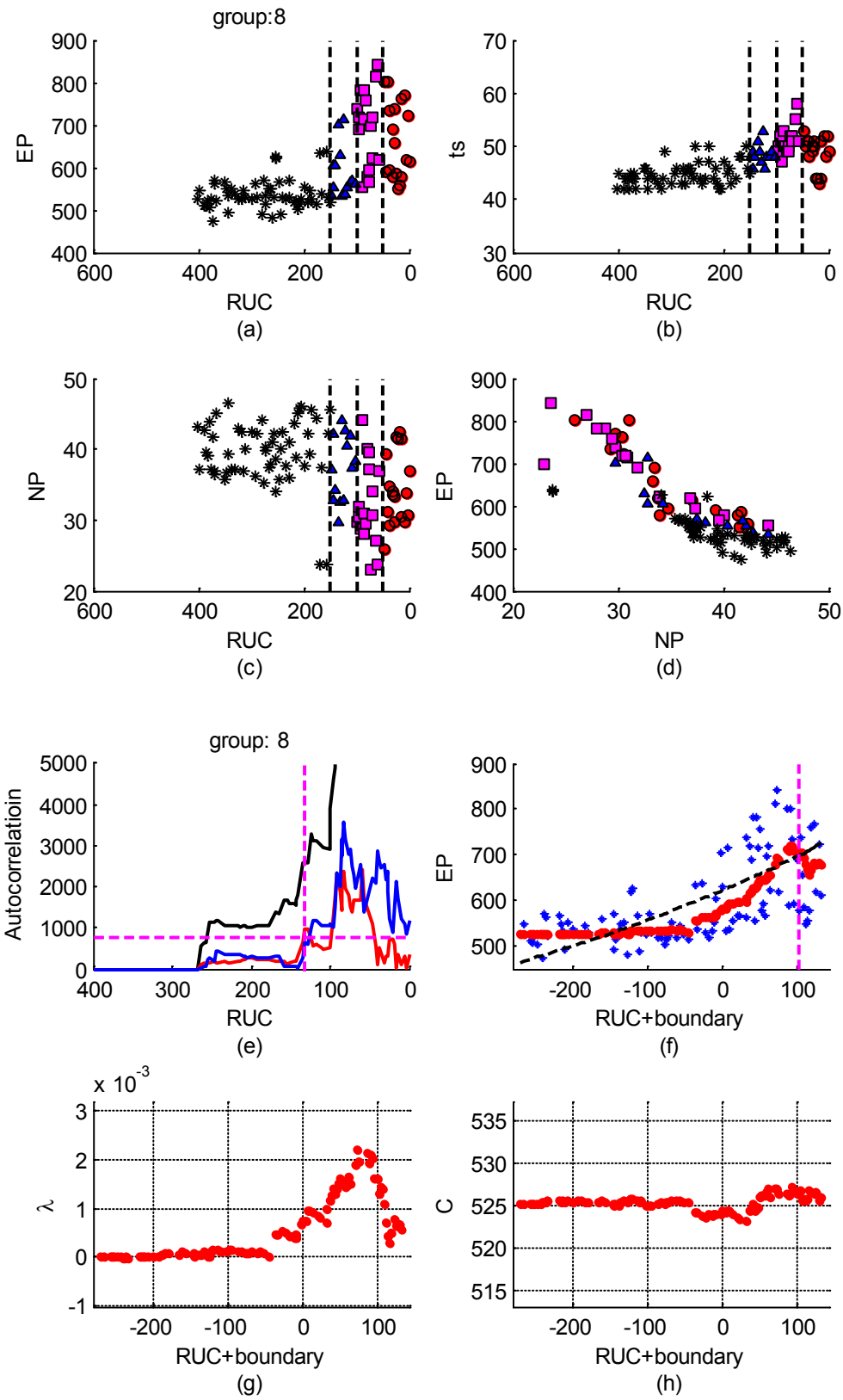

group:3
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 3
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)

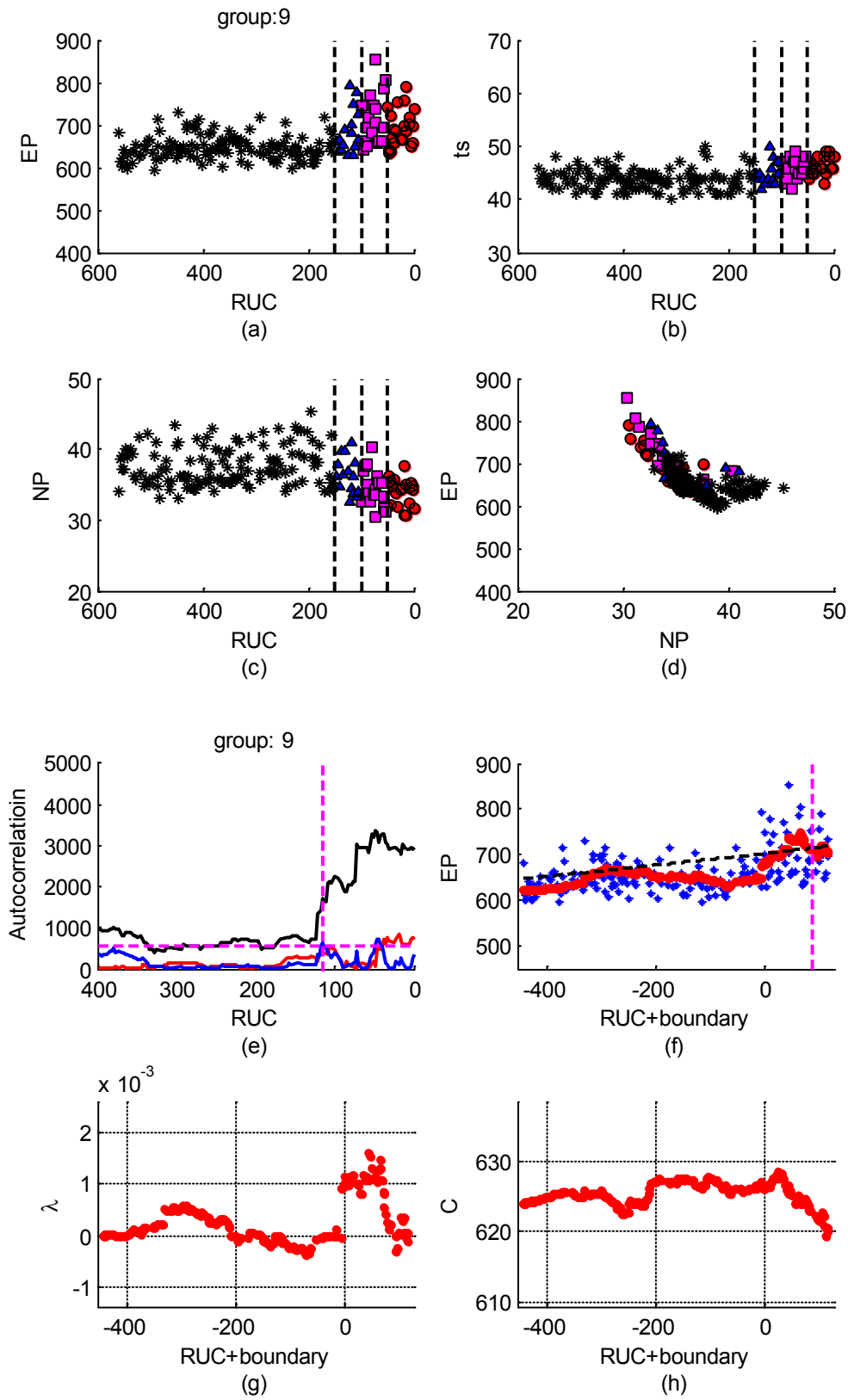

group:4
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 4
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)

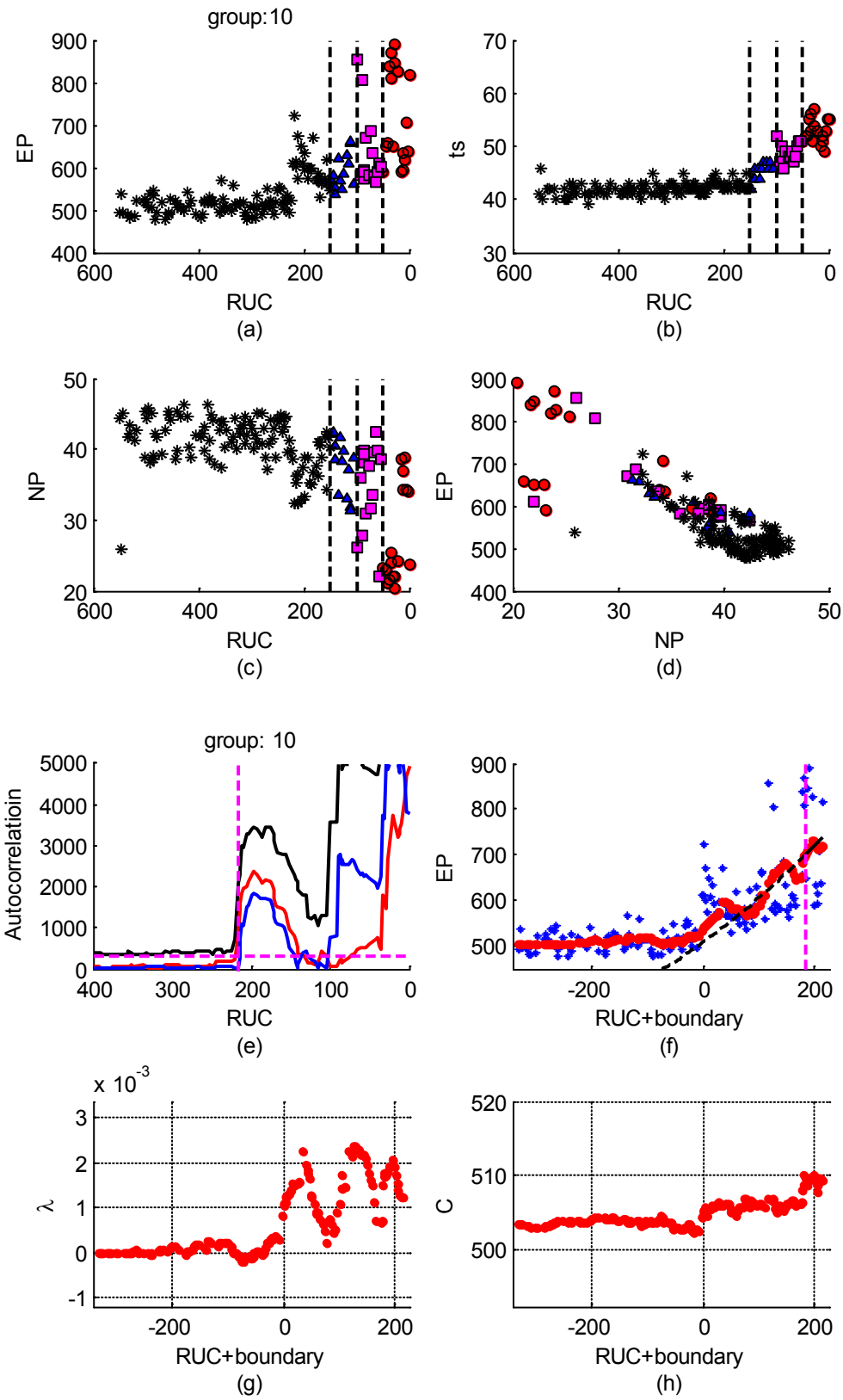

group:5
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 5
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)

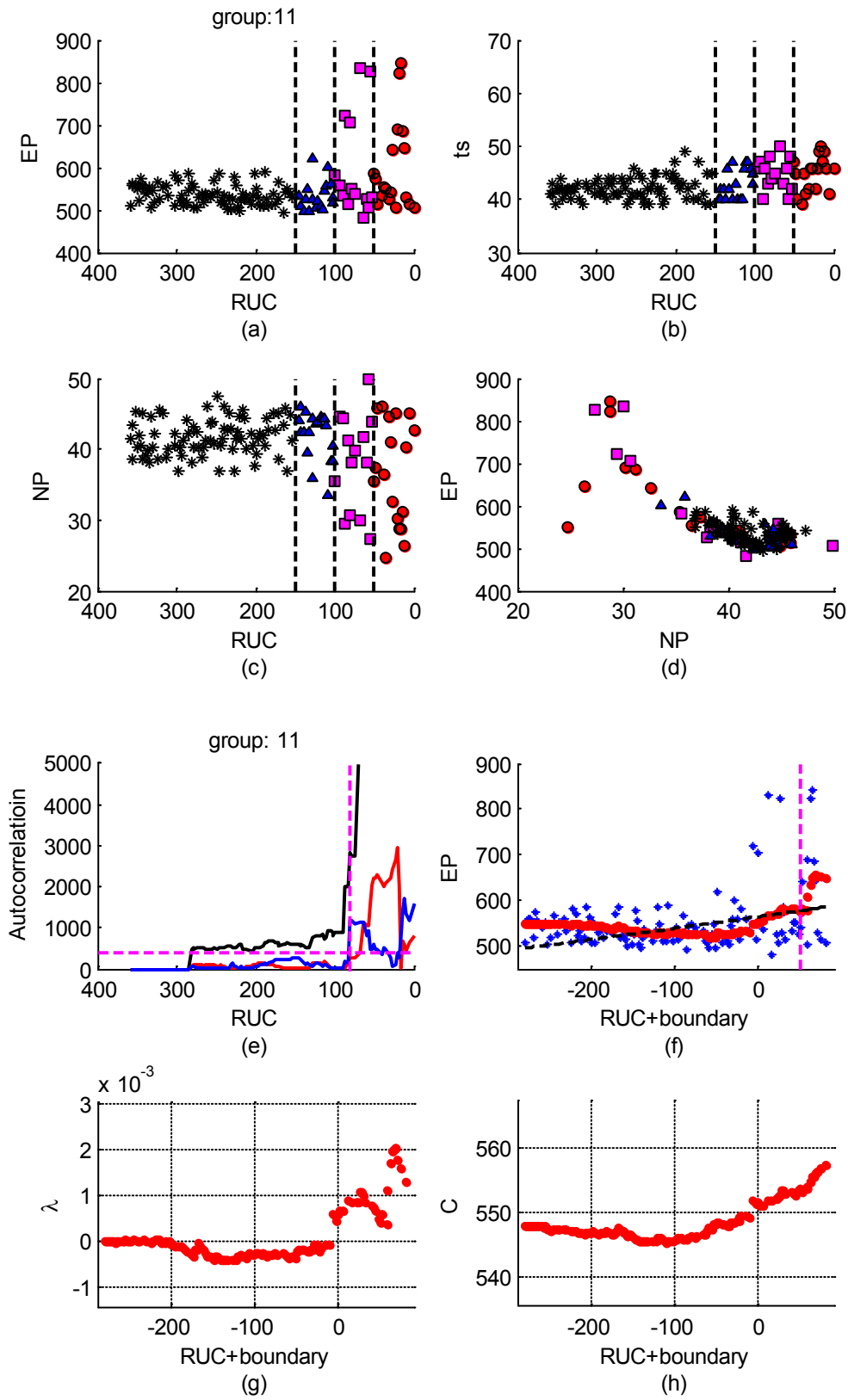

group:6
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 6
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)

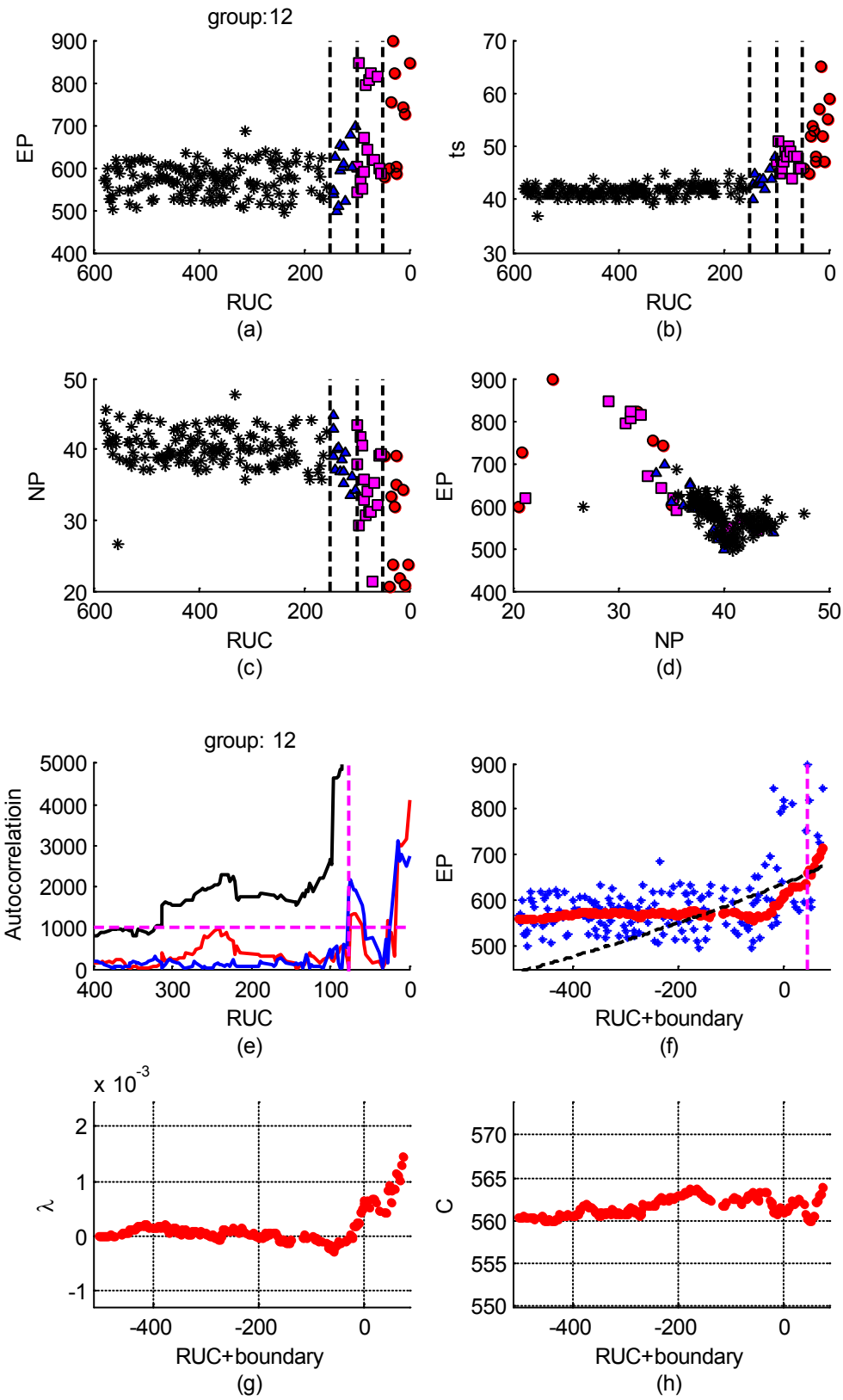

group:7
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 7
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)

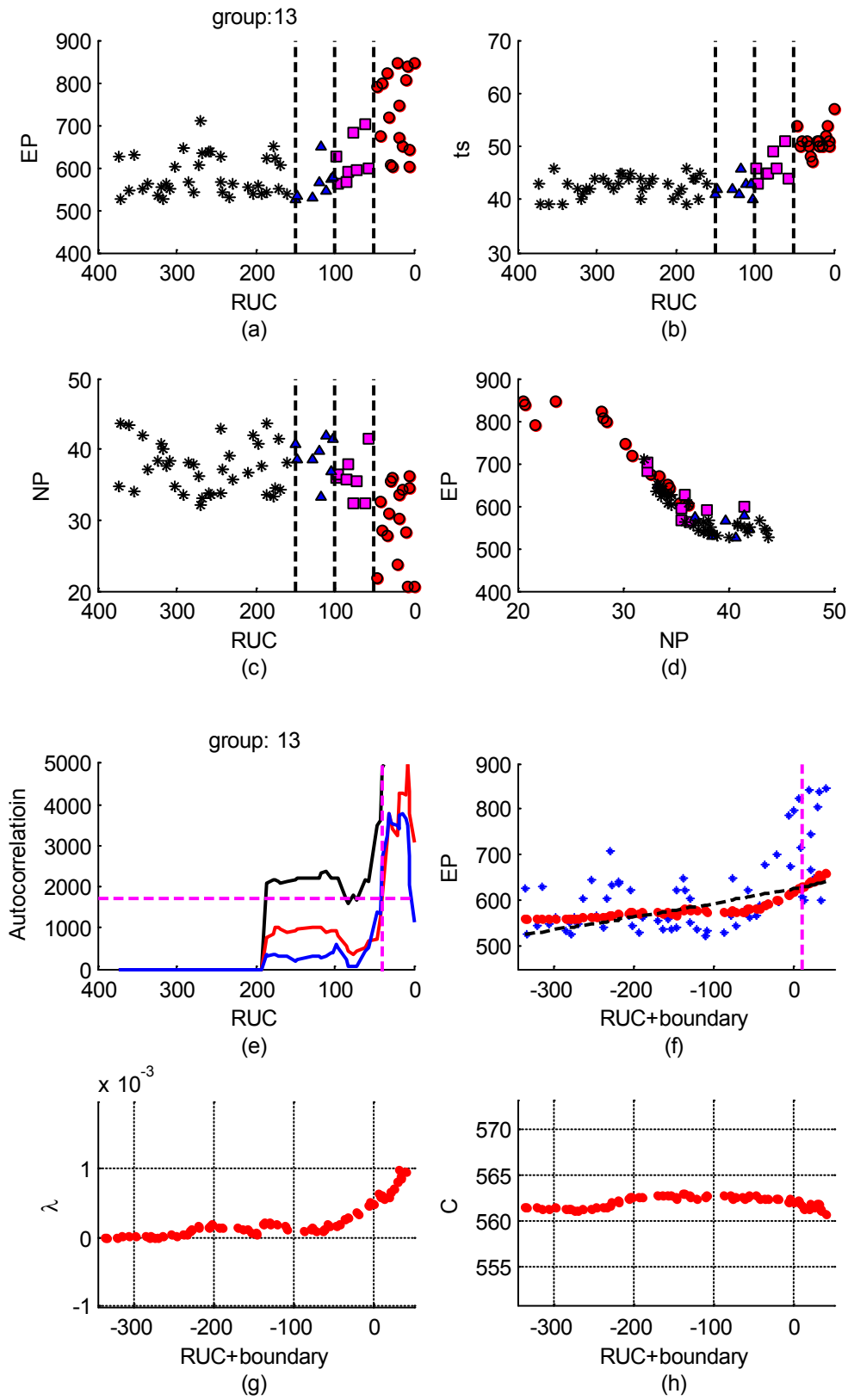

group:8
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 8
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)


group:9
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 9
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)


group:10
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 10
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)


group:11
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 11
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)


group:12
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 12
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)


group:13
EP
RUC
(a)
ts
RUC
(b)
NP
RUC
(c)
EP
NP
(d)
group: 13
Autocorrelatioin
RUC
(e)
EP
RUC+boundary
(f)
x 10^-3
λ
RUC+boundary
(g)
C
RUC+boundary
(h)

Reference

[1] AIRBUS. *A320 Flight crew operating manual*. AIRBUS Inc. 2010.

[2] A. Jardine, D. Lin and D. Banjevic. “A review on machinery diagnostics and prognostics implementing condition-based maintenance.” *Mechanical Systems and Signal Processing*, vol. 20, pp. 1483-1510, 2006.

[3] H. I. Saravanamuttoo, G. F. Rogers and H. Cohen. *Gas turbine theory*. Pearson Education Ltd, 2001.

[4] J. L. Kerrebrock. *Aircraft engines and gas turbines*. vol. 22, Cambridge, MA: MIT press, 1992.

[5] P. P. Walsh and P. Fletcher. *Gas turbine performance*. Wiley, com, 2004.

[6] Rolls-Royce. *The Jet Engine*. Rolls-Royce, 1966.

[7] G. F. Aker and H. I. Saravanamuttoo. "Predicting gas turbine performance degradation due to compressor fouling using computer simulation techniques." *Journal of engineering for gas turbines and power*, vol. 111, pp. 343-350, 1989.

[8] H. I. Saravanamuttoo and A. N. Lakshminarasimha. "Prediction of fouled compressor performance using stage stacking techniques." *Turbomachinery performance deterioration*, pp. 59-66, 1986.

[9] H. Hanachi, J. Liu, A. Banerjee, A. Koul, M. Liang and E. Alavi. "Bladed disc crack diagnostics using blade passage signals." *Measurement Science and Technology*, vol. 23, no. 12, 2012.

[10] K. Kim. "Fault diagnosis and prognosis for fuel supply systems in gas turbine engines." *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol.223, no. 3, pp. 757-768, 2009.

[11] AIRBUS. *A310 Flight Crew Operating Manual*. AIRBUS Inc. 2010.

[12] AIRBUS. *A310 Aircraft Maintenance Manual*. AIRBUS Inc. 2010.

[13] J. Cheng, D. Yu and Y. Yang, "Time-energy density analysis based on wavelet transform," *NDT&E International*, vol.38, pp. 569-572, 2005.

[14] C. M. Harris and A. G. Piersol. *Harris's Shock and Vibration Handbook*. McGraw-Hill, New York, 2002.

[15] R. K. Young. *Wavelets Theory and Its Applications*. Kluwer Academic Publishers, Boston, 1993.

[16] J. Gertler. *Fault Detection and Diagnosis in Engineering Systems*, New York: Marcel Dekker, 1998.

[17] J. Schurmann. *Pattern Recognition: A Unified View of Statistical and Neural Approaches*, Wiley, New York, 1996.

[18] H. Ding, X. Gui and S. Yang. "An approach to state recognition and knowledge-based diagnosis for engines." *Mechanical Systems and Signal Processing 5*, pp. 257-266, 1991.

[19] P. Zhou, H. Li and D. Clelland. "Pattern recognition on diesel engine working conditions by wavelet Kullback-Leibler distance method." *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical*

Engineering Science, vol. 219, pp. 879-887, 2005.

[20] Yu, Jie. "A new fault diagnosis method of multimode processes using Bayesian inference based Gaussian mixture contribution decomposition." *Engineering Applications of Artificial Intelligence*, 2012.

[21] C. Bunks, D. McCarthy and T. Al-Ani. "Condition-based maintenance of machines using hidden Markov models." *Mechanical Systems and Signal Processing*, vol. 14, pp. 597-612, 2000.

[22] M. Ge, R. Du, and Y. Xu. "Hidden Markov model based fault diagnosis for stamping processes." *Mechanical Systems and Signal Processing*, vol. 18, pp. 391 – 408, 2004.

[23] M. T. Hagan, B. D. Howard and H. B. Mark. *Neural network design*. Boston: Pws Pub, 1996.

[24] F. Karray and C. De Silva. *Soft Computing and Intelligent Systems Design: Theory, Tools, and Applications*. New York: Pearson/Addison Wesley, 2004.

[25] L. Kuncheva. "On the equivalence between fuzzy and statistical classifiers." *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems*, vol. 4, no.3, pp. 245-253, 1996.

[26] J. Liu, W. Wang and F. Golnaraghi. "An enhanced diagnostic scheme for bearing condition monitoring. Instrumentation and Measurement." *IEEE Transactions on*, vol. 59, no. 2, pp. 309-321, 2010.

[27] D. C. Baillie and J. Mathew. "Nonlinear model-based fault diagnosis of bearings." *Proceedings of an International Conference on Condition*

Monitoring, Swansea, UK, pp. 241-252, 1994.

[28] C.H. Oppenheimer and K.A. Loparo. “Physically based diagnosis and prognosis of cracked rotor shafts.” *Component and Systems Diagnostics, Prognostics, and Health Management II Bellingham*, vol. 4733, pp. 122-132, 2002.

[29] J. Liu, W. Wang and F. Golnaraghi. "A multi-step predictor with a variable input pattern for system state forecasting." *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1586-1599, 2005.

[30] S. Gupta and A. Ray. "Real-time fatigue life estimation in mechanical structures." *Measurement Science and Technology*, vol. 18, no. 7, pp. 1947-1957, 2007.

[31] S. Yagiz, C. Gokceoglu, E. Sezer and S. Iplikci. "Application of two non-linear prediction tools to the estimation of tunnel boring machine performance." *Engineering Applications of Artificial Intelligence*, vol. 22, no. 4, pp. 808-814, 2009.

[32] C. de Groot and D. Würtz. "Analysis of univariate time series with connectionist nets: A case study of two classical examples." *Neurocomputing*, vol. 3, no. 4, pp. 177-192, 1991.

[33] R. S. Tsay. "Testing and modeling threshold autoregressive processes." *Journal of the American Statistical Association*, vol. 84, no. 405, pp. 231-240, 1989.

[34] T. S. Rao. "On the theory of bilinear time series models." *Journal of the*

Royal Statistical Society. Series B (Methodological), pp. 244-255, 1981.

[35] J. H. Friedman and W. Stuetzle. “Projection pursuit regression.” *Journal of the American Statistical Association*, vol. 76, no. 376 pp. 817-823, 1981.

[36] J. H. Friedman. "Multivariate adaptive regression splines." *The annals of statistics,* pp. 1-67, 1991.

[37] D. Brillinger. "The identification of polynomial systems by means of higher order spectra." *Journal of Sound and Vibration*, vol. 12, no. 3, pp. 301-313, 1970.

[38] A. F. Atiya, S. M. El-Shoura, S. I. Shaheen and M. S. El-Sherif. "A comparison between neural-network forecasting techniques-case study: river flow forecasting." *Neural Networks, IEEE Transactions on*, vol. 10, no. 2, pp. 402-409, 1999.

[39] Y. Liang and L. Xu. "Improving signal prediction performance of neural networks through multiresolution learning approach." *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 36, no.2, pp. 341-352, 2006.

[40] D. Husmeier and J. G. Taylor. *Neural networks for conditional probability estimation: Forecasting beyond point predictions*. Springer-Verlag, 1999.

[41] J. Korbicz ed. *Fault Diagnosis.: Models, Artificial Intelligence, Applications*. Springer, 2004

[42] W. Wang and J. Vrbanek. "An evolving fuzzy predictor for industrial applications." *Fuzzy Systems, IEEE Transactions on*, vol. 16, no. 6, pp.

1439-1449, 2008.

[43] P. W. Tse and D. P. Atherton. "Prediction of machine deterioration using vibration based fault trends and recurrent neural networks." *Journal of vibration and acoustics*, vol. 121, no.3, pp. 355-362, 2009.

[44] D. E. Adams. "Nonlinear damage models for diagnosis and prognosis in structural dynamic systems." *AeroSense 2002*. International Society for Optics and Photonics, 2002.

[45] J. Luo, A. Bixby, K. Pattipati, L. Qiao, M. Kawamoto and S. Chigusa. "An interacting multiple model approach to model-based prognostics." *Systems, Man and Cybernetics, 2003. IEEE International Conference on*. vol. 1. IEEE, 2003.

[46] D. Chelidze and J. P. Cusumano. "A dynamical systems approach to failure prognosis." *TRANSACTIONS-AMERICAN SOCIETY OF MECHANICAL ENGINEERS JOURNAL OF VIBRATION AND ACOUSTICS*, vol. 126, no. 1, pp. 2-8, 2004.

[47] M. Pechtl and J. Rubyca. "A prognostics and health management roadmap for information and electronics-rich systems." *Microelectronics Reliability*, vol. 50, no. 3, pp. 317-323, 2010.

[48] B. Sahar, K. Goebel, S. Poll and J. Christophersen. "Prognostics methods for battery health monitoring using a Bayesian framework." *Instrumentation and Measurement, IEEE Transactions on*, vol. 58, no.2, pp. 291-296, 2009.

[49] A. Doucet, N. De Freitas and N. Gordon. "An introduction to sequential

Monte Carlo methods." *Sequential Monte Carlo methods in practice*. Springer New York, pp. 3-14, 2001.

[50] J. Liu, W. Wang and F Ma. "A regularized auxiliary particle filtering approach for system state estimation and battery life prediction." *Smart Materials and Structures*, vol. 20, no. 7, 075021, 2011

[51] M. E. Orchard and G. J. Vachtsevanos. "A particle filtering-based framework for real-time fault diagnosis and failure prognosis in a turbine engine." *Control & Automation, 2007. MED'07. Mediterranean Conference on*. IEEE, 2007.

[52] S. Letourneau, C. Yang and Z. Liu. "Improving preciseness of time to failure predictions: Application to APU starter." *Prognostics and Health Management, 2008. PHM 2008. International Conference on*. IEEE, 2008.

[53] C. Yang, S. Létourneau, M. Zaluski and E. Scarlett. "Apu FMEA validation and its application to fault identification." Proceedings of ASME. 2010.

[54] A. J. Allan. "Gas turbine parameter corrections." *Journal of Engineering for Gas Turbines and Power*, vol. 121, no.4, 1999.

[55] J. Kurzke. "Model based gas turbine parameter corrections." *Proceedings of the ASME International Gas Turbine And Aeroengine Congress And Exposition*. 2003.

[56] I. D. Coope and C. J. Price. "Positive bases in numerical optimization." *Computational Optimization and Applications*, vol. 21, no. 2, pp. 169-175, 2002.

[57] J. B. Franzini, J. E. Finnemore and R. L. Daugherty. *Fluid mechanics, with engineering applications*. NY: McGraw-Hill, 1977.

[58] A. Papoulis and S. U. Pillai. *Probability, Random Variables, and Stochastic Processes*. Tata McGraw-Hill Education, 2002.

[59] H. Trevor, R. Tibshirani and J. J. H. Friedman. *The elements of statistical learning*. vol. 1. New York: Springer, 2001.

[60] Bárány, Imre and V. Vu. "Central limit theorems for Gaussian polytopes." *The Annals of Probability*, pp. 1593-1621, 2007.

[61] P. G. Broto, G. Moreau and C. Vandycke. "Molecular structures: perception, autocorrelation descriptor and sar studies: system of atomic contributions for the calculation of the n-octanol/water partition coefficients." *European journal of medicinal chemistry*, vol. 19, no. 1, pp. 71-78, 1984.

[62] J. Xia, K. Han and D. Huang. "Sequence-based prediction of protein-protein interactions by means of rotation forest and autocorrelation descriptor." *Protein and Peptide Letters*, vol. 17, no. 1, pp. 137-145, 2010.

[63] J. D. Hamilton. *Time series analysis*. vol. 2. Princeton: Princeton university press, 1994.

[64] R. G. Brown and P. Y. Patrick Hwang. *Introduction to random signals and applied Kalman filtering*. vol. 1, no. 1, New York: John Wiley & Sons, 1992.

[65] J. H. Laning, Jr. and R. H. Battin, *Random processes in Automatic control*. New York: McGraw-Hill, 1956

[66] M. S. Arulampalam, S. Maskell, N. Gordon and T. Clapp. "A tutorial on

particle filters for online nonlinear/non-Gaussian Bayesian tracking." *Signal Processing, IEEE Transactions on*, vol. 50, no. 2, pp.174-188, 2002.

[67] D. Simon. "A comparison of filtering approaches for aircraft engine health estimation." *Aerospace Science and Technology*, vol. 12, no. 4, pp. 276-284, 2008.

[68] N. J. Gordon, D. J. Salmond and A. F. Smith. "Novel approach to nonlinear/non-Gaussian Bayesian state estimation." *IEE Proceedings F (Radar and Signal Processing)*, vol. 140, no. 2, IET Digital Library, 1993.

[69] D. Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. Wiley. com, 2006

[70] A. F. Smith and A. E. Gelfand. "Bayesian statistics without tears: a sampling–resampling perspective." *The American Statistician*, vol. 46, no. 2, pp. 84-88, 1992.