CARACTÉRISATION DE PROCESSUS STOCHASTIQUES À L'OEUVRE DANS LES TURBOMACHINES AÉRONAUTIQUES

Jean-Loup Loyer[1]

[1] Instituto Superior Técnico, *Avenida Rovisco Pais 1, 1049-001 Lisboa*
jean-loup.loyer@tecnico.ulisboa.pt

Résumé. La caractérisation de certains processus stochastiques à l'œuvre dans les moteurs aéronautiques est une tâche complexe qui est déterminante pour comprendre la dégradation des turbomachines, et ce afin d'améliorer la sécurité et de générer des variables explicatives pertinentes pour des modèles statistiques de maintenance prédictive. Toutefois, une telle caractérisation demande d'analyser des centaines de séries temporelles multivariées comprenant chacune des centaines de composantes individuelles présentant des motifs d'évolution complexes. Étant donné qu'une analyse manuelle exhaustive est impossible, il est nécessaire de mettre au point une méthodologie simple, cohérente et reproductible pour identifier les processus stochastiques à l'œuvre dans les moteurs d'avions. Utilisant notamment la représentation graphique des séries temporelles, les fonctions de corrélation croisée et les fonctions d'autocorrélation partielles, nous montrons que de nombreuses interactions entre variables liées au moteur appartiennent à trois grandes catégories de processus (ARMA, bruit blanc, SARIMA) qui correspondent à des types relativement bien définis de phénomènes physiques. Cette méthodologie étant une première tentative de caractérisation de processus stochastiques à l'œuvre dans des moteurs aéronautiques, les défis restant sont présentés en conclusion, avec les voies potentielles pour de futurs travaux de recherche.

Mots-clés. Processus stochastiques, analyse de séries temporelles multivariées, dégradation de turbomachines, maintenance prédictive

Abstract. Characterizing the stochastic processes at stake in jet engines is a complex task that is important to understand the deterioration of turbomachines so as to improve safety and generate relevant predictors for statistical models of predictive maintenance. However, it requires the analysis of hundreds of multivariate time series comprising hundreds of individual components displaying complex evolution patterns. Since a manual thorough analysis is not possible, it is necessary to design a simple, coherent and reproducible methodology to identify the stochastic processes at stake in jet engines. Using notably time series plots, cross-correlation functions and partial auto-correlation function, we show that many interactions between engine-related variables fall into three large categories of processes (ARMA, white noise and SARIMA) corresponding to relatively well-defined types of physical phenomena. This methodology being a first attempt at characterizing stochastic processes in jet engines, remaining challenges are finally presented, along with potential avenues for future research.

Keywords. Stochastic process, multivariate time series analysis, turbomachines deterioration, predictive maintenance

1 Introduction to the problem

Jet engines are complex machines whose performance and reliability are strongly influenced by internal thermodynamical and mechanical phenomena as well as external environment-related events. Although a turbomachine is mostly driven by the deterministic laws of physics, it is also partly subject to complex random phenomena, which are often analyzed less thoroughly by aerospace engineers. Nonetheless, understanding the stochastic processes at stake in jet engines might contribute to improve their performance (i.e. fuel consumption), probe potential faults and predict maintenance needs as explained by Heng, Zhang and Mathew (2009).

However, the problem involves potentially hundreds of internal and external variables recorded on thousands of engines: a manual characterization – and a fortiori classification - of the stochastic processes are thus not feasible. Instead, it is worthwhile to develop a method that would (semi)automatically identify the nature of the stochastic process and estimate its main parameters.

Identifying the characteristics of the stochastic processes at the system level is interesting from an academic perspective, as the contributions to the field are still scarce and tend to focus on a few variables at the component level. For instance, Zhan, Makis and Jardine (2003) have focused on modelling gearbox faults based on Kalman filtering of multivariate vibration signals while Baillie and Mathew (1996) have used autoregressive models for diagnosing rolling elements bearings. It also has an industrial interest since it can help understanding the correlation structure in the dataset and extract relevant features for later statistical modelling of maintenance costs.

After introducing the structure of the dataset, the communication will present the methodology for identifying the nature of the stochastic processes, followed by a first attempt at identifying categories of processes. The discussion will finally cover remaining challenges and future works.

2 Structure of the dataset

The dataset consists of multivariate times series whose components can be divided into two sets of variables: 1) internal parameters of thermodynamical (temperature, pressure, shaft rotation speeds, fuel flows…) or mechanical (vibration levels, oil temperature and pressure…) nature, 2) external parameters related to the atmospheric environment the engine is operating in (air temperature & pressure, humidity, altitude, Mach number…). While an engine is operating, such parameters are acquired at the frequency of 1Hz by high-quality redundant sensors located in several stations inside and outside the engine. The data are processed by the Engine Electronic Controller (EEC) to control the engine while “snapshots” are extracted at key moments of each flight (maximum take-off power, mid-climb, cruise…) to be passed to the engine Health Monitoring Unit (HMU) and subsequently sent to the ground in near-real-time for engineering analysis. The quality of the data is excellent given the high requirements of the data acquisition system: the sensors have stringent technical specifications while the data processing algorithms have been carefully designed by highly skilled engineers specialized in control systems. Any deviation in the data is rigorously identified and analyzed in near-real-time so that faulty electronic equipment can be changed within days. Thus, the remaining outliers are so few that they don’t influence significantly the analysis of the time series (Figure 1).

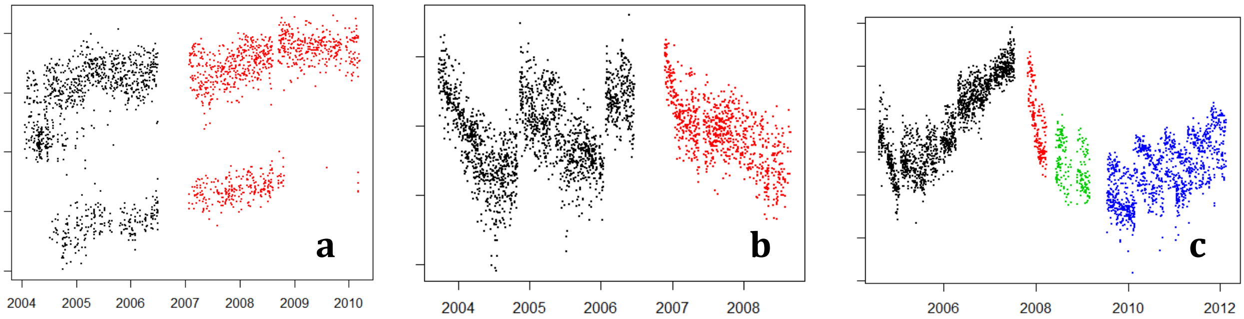


Figure 1 - Examples of representative time series in the dataset (fictitious data)

The data analyzed in this communication corresponds to the snapshots taken at maximum take-off, which is the most damaging phase of a flight for a jet engine. A “data point” is considered as one measurement:

- for each of the hundreds of internal and external variables aforementioned
- for each flight over a 10 year period (there is approximately one flight per day)
- for the hundreds of engines in the fleet currently in service

Therefore, the dataset consists in a set of hundreds of multivariate time series each comprising hundreds of components containing each a few hundred points. We are interested not only in characterizing the stochastic process for individual components but also between the components of a multivariate time series. Fictitious[1] yet representative examples (Figure 1) show a large variety of complex patterns in the times series, including in some cases (non)linear trends and/or seasonality and/or inconstant variance and/or varying autocorrelation structure over time and/or steps/gaps: some engine parameters evolve almost linearly (red segment of curve b) while other exhibit nonlinear trends (black segment in curve c), erratic patterns (black segment of curve b) or clusters of values (curve a). Some engine parameters exhibit more variability (curves a, b) than others (curves c). Finally, some curves present many discontinuities either as steps (curves b, c) or gaps in the data (curves c). Given the potential combinations of patterns, the diversity of the stochastic processes is likely to be high. We present in the rest of the communication a systematic method to identify them more rigorously an semiautomatically.

3 Methodology for characterizing the stochastic processes

Characterization of stochastic processes is a complex intellectual task involving model identification, parameter estimation and model diagnosis, according to Box and Jenkins (1994) three-step generic methodology (Figure 2).

[1] To respect confidentiality agreement with the industrial partner, fictitious data has been simulated, with changes in dates in the X-axis and absence of variable name in the Y-axis. Each colored segment corresponds to the interval of the engine operates between two maintenance visits and is considered as an individual component of a multivariate time series in the dataset.

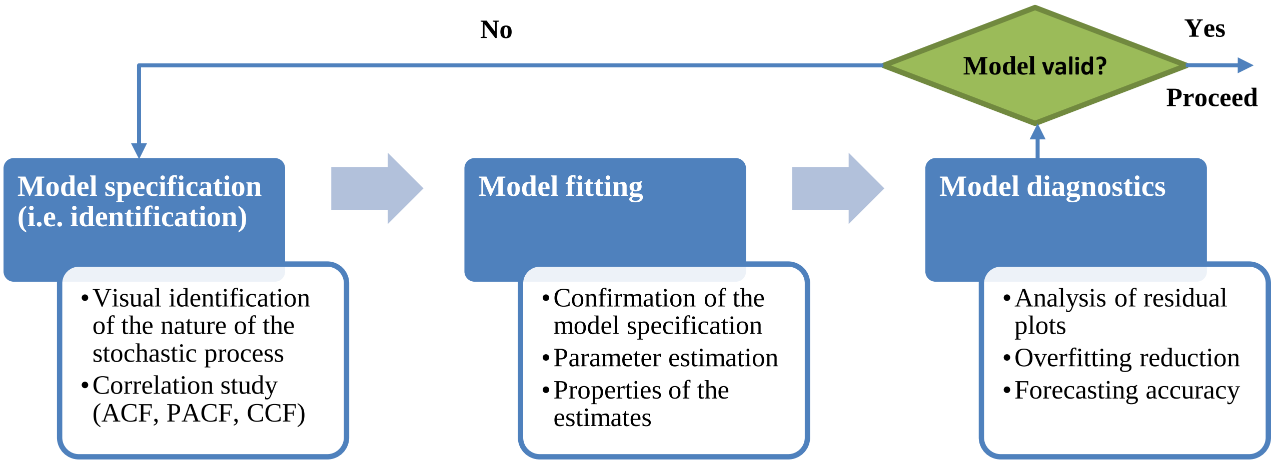


Figure 2 - Bok and Jenkins three-step method for the characterization of time series and stochastic processes

This communication focuses on the first phase of the methodology and presents early results on the model fitting. In practice, identifying the stochastic processes in a straightforward way could be done by plotting, for hundreds of engines and dozens of pairs of variables, the original time series (Figure 1) complemented by lag plots for various values of lags. However, the number of plots would amount to millions and soon become overwhelming: instead we compute and plot, for hundreds of engines in the fleet[2], the Cross-Correlation Function (CCF) between all pairs of variables in the dataset for lags comprised between -300 and +300, spanning approximately plus or minus one full year of engine service (Figure 3). The CCF plots are then complemented by plots of the Partial Auto-Correlation Functions (PACF) to refine the models identification and have an initial idea of the parameters for autoregressive models (not displayed in this communication).

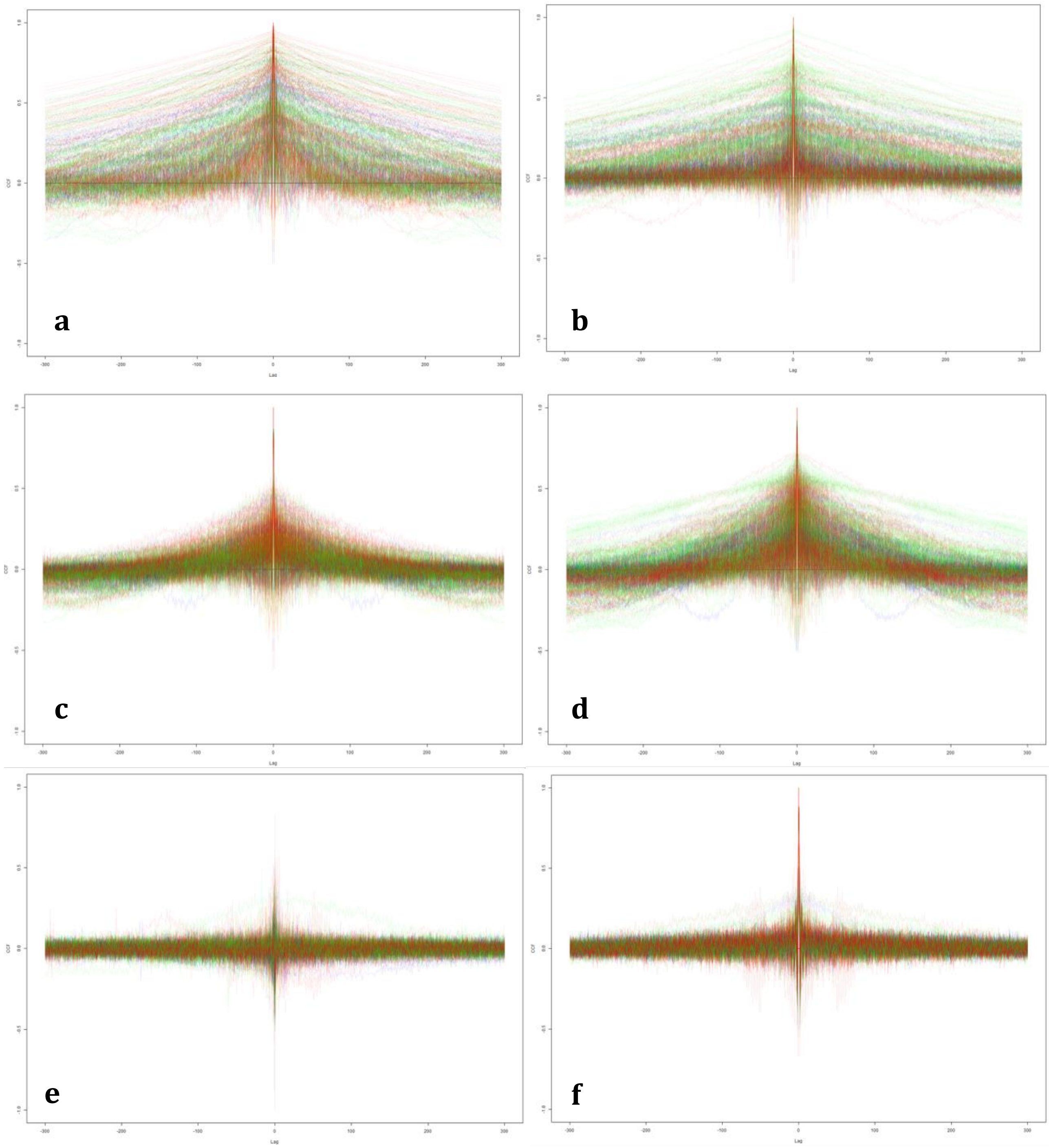
The visual analysis of the original time series, the ACF, PACF and CCF allows one to identify the nature of the stochastic process amongst a broad set of potential candidates. Many times series in the dataset can be modelled, before or after transformation, by several of such models concurrently. The estimation of the parameters is another complex iterative step that is out of the scope of the communication, although likely intervals for some parameters will be specified for key pairs of engine parameters.

4 Results: a high diversity of stochastic processes in jet engines

The analysis of time series demonstrates the high diversity in the stochastic processes involved in jet engines, as illustrated by the hundreds of superimposed CCFs for selected pairs of variables (Figure 3).

---

[2]Each colored curve of the plot in Figure 3 corresponds to a “period” in the life of a given engine, namely an interval between two maintenance visits. The first period covers the time between Entry Into Service (EIS) and the first Maintenance Shop Visit (MSV), usually a light check: it corresponds to a new engine that has not yet deteriorated but may still present early random defects due to design, manufacturing or handling problems. The second period covers the interval of time between the first and second MSVs, where more important reparations are usually performed. The third period between the second and third MSVs corresponds to a “mature” engine with heavy (resp. minor) repairs if the previous MSV involved light (resp. heavy) maintenance.


a
b
c
d
e
f

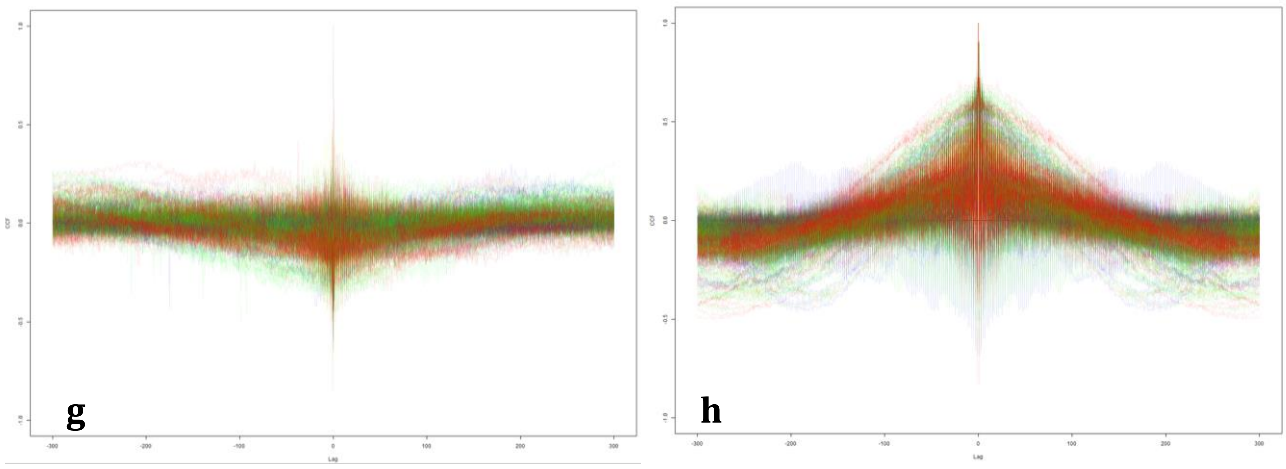


Figure 3 - Illustration of the diversity of the stochastic processes in turbomachines through cross-correlation functions

After a visual review of hundreds of such plots, we identified three main categories of stochastic processes:

1. The evolution in time of some (interactions of) variables can be approximated by simple processes such as moving average (MA_p), autoregressive (AR_q), autoregressive moving average ($ARMA_{p,q}$) with p and q equal to 1 or 2 (four top curves a-d). Such processes often correspond in fact to Auto-Correlated Functions (ACF), namely the CCF of a variable with itself, as can be seen by a correlation equal to 1 at zero lag and by the symmetry of the curves. Such models have already been used for other applications, notably by Baillie and Mathew (1996)
2. Other stochastic processes can be considered as white noise processes (curves e-f) and often correspond to CCF of specific engine parameters (e.g. a specific temperature in a given station of the engine) against engine parameters at the system level (e.g. engine thrust …).
3. A third category of stochastic processes are seasonal autoregressive integrated moving average ($SARIMA_{p,d,q}$) models with more complex intertwined patterns (curves g-h), which can often be attributed to CCF of an internal engine parameter against an environmental parameter (e.g. external air temperature).

Many stochastic processes of engines having operated in typical conditions can be relatively clearly attributed to one of those categories, whereas it is usually not as clear for engines that have undergone less nominal operating conditions (installation in aircraft of very different airliners, damage by foreign object…).

5 Conclusion and discussion

To our knowledge, this communication is the first attempt at characterizing stochastic processes at stake in jet engines, at the system level and based on large volumes of actual data. This introductory analysis demonstrates that randomness involved in such a complex machine shares common characteristics with phenomena as diverse as particles movements in physics or stock evolution in finance, according to Bhattacharya and Waymire (1990). Moreover, although the variety of the stochastic processes is high, it is possible to distinguish a few categories corresponding to specific types of physical phenomena.

However, the characterization of stochastic processes in turbomachines is not only of abstract academic interest: the results of the analysis can indeed be translated into engineering insights, with a direct practical use for the business analysts of manufacturers and airliners. In particular, identification and parameter estimation of stochastic processes allows one to classify engines into relevant categories, isolate engine-specific effects from environmental effects, detect anomalies and focusing on problematic assets by identifying outliers via the characterization of the stochastic processes. Such engineering-based analysis directly leads to improved customer service and higher profit through better forecasting of airworthiness-related issues, aftersales needs and maintenance planning.

Nonetheless, the analysis presented in this communication is a mere introduction to stochastic processes in jet engines, whose understanding and modelling for thousands of engines and potentially hundreds of variables remains a difficult challenge. The stochastic processes of the variables and their interactions have to be further studied so as to reduce the dimensionality of the problem: transformation of the original time series to obtain time series with better properties for modeling such as stationarity (via differencing, logarithm, square root or Box-Cox transformation…), extension to multivariate models, analysis of cointegration, construction of regression models... The exact nature and parameters of the stochastic processes would also have to be further assessed by more sophisticated techniques in order to identify meaningful patterns and clusters of time series, with the ultimate objective to extract relevant features for subsequent predictive models of maintenance.

Bibliographie

[1] Baillie, D.C. et Mathew, J. (1996), A comparison of autoregressive modeling techniques for fault diagnosis of rolling element bearings, *Mechanical Systems and Signal Processing*, 10, 1–17

[2] Bhattacharya, R.N. et Waymire, E.C. (1990), *Stochastic processes with applications*, Wiley, New York.

[3] Box, G. E. P., Jenkins, G. M. et Reinsel, G. C. (1994), *Time Series Analysis, Forecasting and Control, 2nd ed*. Prentice-Hall, New York.

[4] Heng, A., Zhang, S., Tan, A.C.C et Mathew, J. (2009), Rotating machinery prognostics: State of the art, challenges and opportunities, *Mechanical Systems and Signal Processing*, 23(3), 724-739

[5] Zhan, Y., Makis, V., et Jardine, A.K.S. (2003), Adaptive model for vibration monitoring of rotating machinery subject to random deterioration, *Journal of Quality in Maintenance Engineering*, 9(4), 351-375