DOI: 10.13009/EUCASS2017-417

Promising HMS approaches for liquid rocket engines

Alessandra Iannetti, Julien Marzat**, Hélène Piet-Lahanier**, Camille Sarotte **, Gerard Ordonneau ***

**CNES*
52 rue J. Hillairet, Paris, France

***ONERA The French Aerospace Lab*
Chemin de la Hunière, Palaiseau, France

Abstract

The improvement of diagnosis systems for liquid rocket engines is an important step towards more competitive and cost effective propulsion systems. The classical monitoring approaches used at bench or in flight are based on fixed thresholds redlines over critical parameters identified from past anomalies and engine behavior expertise. Decades of testing and launching have shown the limitation of this approach especially when the number of firings and launches increases or when reusable systems are concerned. Today to follow with the new ambitions of cutting the cost to access to space for both expendable and reusable systems, new diagnosis algorithms have to be considered and tested in order to provide new solutions. After reviewing the main needs of a future innovative propulsion system in terms of possible monitoring and diagnosis needs, we propose a review of the most promising approaches in HMS based on recent research works.

1. Introduction

Rocket engines are high energy complex systems. Producing up to 12000 kN of thrust (corresponding to 1000 kg/s), thanks to the conversion of the high energy of the combustion product into high speed ejected mass flows. The energy levels can reach up to several GW of power all in a very confined space. Rocket engines have to withstand harsh thermal and mechanical environment and to provide a high level of reliability: unexpected events can lead to catastrophic anomalies in a very short time (10 to 30 ms).
This is why since the very early years of rocket engine testing and flight the issue of detecting anomalies early has been a fundamental key to success. The typical detection system is based on fixed thresholds applied on critical sensors outputs. The choice of the sensors, their location and the values of the allowed thresholds are based on the failure modes analysis and on the known anomalies.
This classical approach is indeed the simplest possible solution and while it has the main advantage of being simple to implement with no computational needs, it has several disadvantages such as:

- The need to account for uncertainties in the thresholds selection leads to detection delay
- There is no knowledge of the system status while it lies in the allowed thresholds
- False alarm may occur due to errors in thresholds implementation
- Setting and validating the fixed thresholds, which are different for each operating point, induces an important workload
- The need to have sensor redundancy

Classical redlines approaches have mainly been deployed in the early years of rocket engine development due to the limited computational resources on bench or on board. Today the situation is completely new and with powerful computers at hand and new methods to analyze data in real time or differed time; this opens new perspectives in health management systems for rockets. To overcome the limitation of the current diagnosis approaches, multiple ways exist that are based on innovative algorithms to process the data coming from the existing sensors. The objective of such an innovative system will be to reduce detection delays, increase knowledge of the status of the system and reduce the workload related to specific thresholds setting and anomalies analysis. This will bring a new flexibility to bench testing, open new ways for flight recovery in case of anomalies; contribute to maintenance strategies for reusable rocket engines, and last but not least reduce the need for sensor redundancy.

DOI: 10.13009/EUCASS2017-417

2. HMS approaches

When talking about diagnosis and prognosis for rocket engines, we refer to the overall system with the term HMS, Health Monitoring System. The main functions are: detection, diagnosis and prognosis.
This article will focus on the first two fundamental functions, detection and diagnosis, with an overview of the main approaches for a successful data treatment.

There are two main families of approaches:

- Model based [43]
- Data driven [44]

Data driven methods are extensively used in most industrial domains nowadays. All these techniques can be referred to as Big Data approaches. The more data available, the more accurate the output of these algorithms would be.
In the field of rocket engines, data amount can be limited: for Ariane 5 around 5 to 6 launches per year are performed, which is a very different situation from airplanes with hours and hours of engine functioning.
The main limitation of data driven approaches such as neural network identification is that they need to be trained and qualified with a somehow fair enough amount of data.
Another approach is the model-based one. Its reliability is dependent on the knowledge of the process and its model.
In many industrial domains the difficulty is that the model of the process is too complex. An example is the modeling of the core nuclear reactions and the interactions with the fluid processes within the nuclear reactor.
In rocket engines, simulation and models have been a complex business too for a long time, but after decades of work all main industrial actors make an extensive use of simulation tools at all steps of the design of the engine. Many complex problems that could not be modeled 30 or 40 years ago have now dedicated codes and have been validated through testing and flights [7] .
This modeling capability represents a specific potential for rocket engines compared to other industrial domains with respect to the development of model based diagnosis methods.
A set of data of reduced size and increasingly efficient modeling capabilities are an asset for the choice of model based HMS strategies.

3. Model based algorithms

Model based diagnosis algorithms use equations of the process to provide observers and indicators. The equations are usually rearranged in the form of a state vector and measured or non-measured input and outputs.
The system state is then estimated generally by minimizing the error between the model prediction and the sensor output via an optimization process.
This error is used for detection and diagnosis purposes, and is usually called a diagnosis residual.
The main techniques to achieve model based diagnosis are:
- Parameter identification via least square minimization [1]
- State observers and Kalman filters [8]

In the first case the objective is to identify the constant parameter of the equation while in state observers the equations are used to predict the behavior of the state variables.

3.1 Examples of parameter identification algorithms

Parameter identification algorithms are designed to determine non measurable parameters of a mathematical model representing the relationship between input and output of a system.
The principle is to detect anomalies based on abnormal variations of the estimated parameters.
By defining a generic process model

$$Y = f(U, \theta, X) \qquad (1)$$

with Y and U vectors of the measured input and output, X is the state vector, θ is the parameter vector to be identified. These parameters can be expressed via more or less complex physical correlations. The faults with an impact on these coefficients will have an impact on their values and will be detected. The function f can be based on polynomial expression for a static process, or on differential equations in the dynamic case. For a static process the

DOI: 10.13009/EUCASS2017-417

parameters are estimated by minimizing the prediction error between estimated outputs and those predicted by the model with the available signals and by applying the optimum criterion of the least squares which is equivalent to the maximum likelihood criterion for the linear case [1].
For a model linear in the parameters, the system of equations can be written as:

$$Y_N = f(U,\theta,X) = H_N\theta + e \qquad (2)$$

Where N corresponds to the number of model outputs, H_N represents the N algebraic equations of the type *f(U,X)* and *e* is the model error which can be represented via a cost function such as:

$$V = \sum_{k=1}^{N} e_k^2 = \|Y_N - H_N\theta\|_2^2 \qquad (3)$$

where the convex optimum is given by $\frac{dV}{dt} = 0$.

The least square estimation of the vector of the model parameters θ is given by the following expression:

$$\hat{\theta} = \arg\min\|Y_N - H_N\theta\|_2^2 \qquad (4)$$

Which can be expressed as:

$$\hat{\theta} = (H_N^T H_N)^{-1} H_N^T Y_N \qquad (5)$$

The value of V computed with this value of parameters is thus minimal.
This approach gives good results for the identification of parameters constant with time or subject to slow evolutions [2][3][4], in particular in steady state operation. Demonstration on real time firing results showed that parameter identification allows good detection rates but the algorithm can be sometimes too sensitive to perturbations and results in high false alarms when noise levels reach the minimum detectable levels [5].

3.2 Examples of Kalman filter and observer algorithms

Kalman filtering [8] approaches allow to take into account stochastic behavior of the process such as noise and modeling errors [9], unlike Luenberger observers [10][11]which are deterministic.
The Kalman filter is designed to estimate the state of a dynamic linear system in presence of additive white Gaussian noise [12]. Given a dynamic model, it is possible to develop a Kalman filter and estimate online the state of a system if it is observable. The detection and diagnosis can then be realized via the analysis of the prediction error given by the difference between the observed status and its prediction.
Given a state representation of a physical process, at discrete time k, with a state noise to include modeling uncertainties, and observation noise (sensor noise):

$$\begin{cases} X_k = f_k(X_{k-1}, U_{k-1}) + w_k \\ Z_k = C_k(X_k) + v_k \end{cases} \qquad (6)$$

$X_k \in \Re^n$, $Z_k \in \Re^m$ et $U_k \in \Re^l$ representing respectively the vectors of state, observation, command or inputs and f_k et C_k are fonctions describing the dynamics of the state vector or the output vector and transformation between the states and the measurements Z_k, w_k and v_k represent the noise on the state and the observation [8], supposed here as white and mutually non correlated i.e. $E\{w_k v_k^T\} = 0$, with zero average and of covariance matrix $Q_{bk} = E\{w_k w_k^T\}$ and $R_k = E\{v_k v_k^T\}$ given as known. In this case the inputs U_k include also additive noise of zero average and of covariance matrix Q_{uk}.
When a model of all the possible sources of uncertainty (w and v) is available, generally under the form of stochastic vectors with known densities, Kalman filtering methods allow to generate residuals sensitive to faults, with different versions to deal with linear or nonlinear models.
The Kalman filter approach provides an estimate $\hat{X}_k$ of the state of the system such that the variance of the estimation error is minimized, i.e.:

$$\hat{X} = \arg\min E\{\hat{Y}_k \hat{Y}_k^T | Z_k\} \qquad (7)$$

where $\hat{Y}_k = X_k - \hat{X}_{k|k}$ is the estimation error or Kalman residual.

DOI: 10.13009/EUCASS2017-417

With the extended version of the filter (EKF) [14][15][16][17], it is also possible to extend the method to nonlinear f_k and C_k functions thanks to a first-order Taylor development of the model around the current state estimation $\hat{X}_{k|k}$ [18][19]. The state estimated value and its covariance matrix are obtained thanks to the recursive application of the cycle of prediction and correction [20]. Some examples of application of a simple EKF are provided in [5]. In this reference the EKF is used for the detection of the cooling system of a test bench. The results obtained on real data sets of bench firing tests show very good detection rates and good filtering of noise from the sensors.
The efficiency of the Kalman filter depends on the quality of the linearization of the functions f around the estimation of the current state $\hat{X}_k$. As long as the current estimation is close to the real state of the system linearization is valid but if the system derives from too much, the linearization may become less reliable [21] and divergence of the filter can occur. The implementation of the method may also be complex due to the need of computation of the Jacobian matrix at each sampling step. If the convergence speed for the parameter estimates to the real values is slower than that of the state variables the numerical robustness can be affected [22].
Extensions of Kalman filter exist such as the Extended Kalman Filter - EKF to overcome these limitations. Some additional examples are [23],: Second order EKF (SOEKF) which calculates the Hessian matrix and limit the risks of divergence of the estimator due to first order Taylor development errors, the unscented Kalman filters (UKF) [24]□[24] which, under the hypothesis of state variables corrupted by white Gaussian noise, allows to choose in a deterministic way the observations to use so as to ensure an average value and a covariance always close to the reality for the estimation. Another alternative is to use different techniques of polynomial approximation of the nonlinear functions, such filters are presented in [19] and applied in [26] for the detection of oscillatory faults in the control loop of the control surfaces of A380.
Other promising algorithms are Unknown Input Observers (UIO). This approach can be used to estimate the state of a system, treating the failure as an unknown input [45]. The Unknown Input Observers (UIO) can be applied to linear systems with both known and unknown inputs [26][27]. The motivation for developing these techniques comes from the limitation of traditional observers and the need to identify unknown disturbances. To tackle this problem, unknown inputs may be modeled by the unknown response of a suitable chosen dynamical system. A simple method of UIO consists in designing a full order observer for linear systems with unknown inputs and using a coordinate system transformation to decouple the disturbance effect on the system and obtain a homogenous equation of the dynamic observer error, as in [28]. The coordinate system transformation modifies the system in such a way that a standard observer design can be applied and then be solved using the Luenberger theory [29]. A simultaneous estimation of the system states and the unknown inputs for linear systems when the so-called observer matching condition is not satisfied has been proposed by [30]. An auxiliary output vector is introduced so that the observer matching condition is satisfied and is used as the new system outputs to asymptotically estimate the system states without suffering the influence of the unknown inputs. The problem of designing UIO for non-linear systems has also been studied with extended UIO consisting on linearization techniques for non-linear discrete-time systems with unknown inputs [31]. The structure of this observer is very similar to that of the Kalman filter in the EKF form. As presented in [22], the state estimation results can be used in two ways:

1. to directly use the state estimation residuals to perform detection via comparison of the predicted state to its measure
2. to follow non nominal variation of non-measurable internal parameters of the model (states and/or parameters), for this the state estimation can be used together with some consistency tests in the parameter space or the state space [13].

Application of EKF for detection on real data has been performed in [5]. Very good results were obtained on the firing tests of a cryogenic bench: with respect to the parameter identification, Kalman filtering gives much better results in terms of false alarm. This approach allows to filter much better the noise behavior but on the other side is less sensitive to slow degradation of amplitude comparable to the nominal noise.
Application of UIO and Kalman filtering is currently being tested with promising results on rocket engines: for example, critical non measured variables such as the mass flows can be easily predicted online based on the functional model and standard pressure and temperature sensors.

3.3 Residual analysis methods

A fault detection system is usually decomposed in a residual generator [32] to quantify a change in the process and a residual evaluator to assess whether the estimated variation is nominal or not [46].
The residual estimation function can be obtained using different algorithms that consider the sequence of acquired residuals as independent random variables with a probability density depending upon only one scalar parameter [32]. One common tool to design on-line change detection is the log-likelihood ratio: in this approach a change in the

monitored residual is reflected as a change in the sign of the mean value of the likelihood ratio, which can be interpreted as the property of detectability of the change [47]. The detectability of a change can also be defined with the help of the Kullback information. Several simple and well-known algorithms for residual analysis have been developed such as the Shewhart control chart, the geometric moving average control chart, the finite moving average control charts, filtered derivative algorithms and CUSUM algorithms, Bayes type's algorithms and generalized likelihood ratio (GLR) algorithms, for more details see [38]. The CUSUM control chart has been widely used to detect mean shifts. The CUSUM algorithm is equivalent to a repeated sequential probability ratio test, the SPRT defining a decision rule and a stopping time with respect to the Neyman-Pearson rule. The decision rule can be seen as the integration of the observations over a sliding window with random size, see e.g. [39]. To detect a positive or a negative change a two sided CUSUM algorithm has been proposed by [40], one side algorithm to detect an increase, another to detect a decrease. Since those algorithms need prior knowledge on the change, relatively simple and computationally inexpensive methods have been developed to estimate the fault size such as the exponentially moving average algorithm developed by [41]. Those methods are referred to adaptive CUSUM (ACUSUM) charts where the estimated shift is updated by the exponentially weighted moving average (EWMA) method; the reference value and the control limit are dynamically adjusted. The numerical results show that the ACUSUM chart can detect a small shift faster but for a large shift it may require more observations. In [42] a modified ACUSUM chart has been proposed using a linear weight on the chart statistic.
Different approaches of automatic thresholds coupled to the CUSUM test were tested in [6] showing good results with both of the approaches for detection used.

4. Conclusions

Thanks to new computational capabilities it is today possible to foresee the use of innovative approaches for real time detection and diagnosis of liquid rocket engines. A great potential is represented by model based approaches coupled with mathematical methods reducing the sensitivity of the detection process to modeling uncertainties and sensors noise. Two examples of promising approaches are Parameter identification and state observers (Kalman filtering in particular), and some combination of the two, because they allow an efficient treatment of the data in real time and because of the availability of reliable prediction models for the majority of the engine processes. Statistical tests such as CUSUM or likelihood ratio methods are then used to determine the detection flags on the basis of the generated residuals. These fault diagnosis methods have been widely used in many industrial domains and are now mature enough to be integrated in the development cycles of new rocket engines, in particular in the perspective of reusable launchers.

5.Acknowledgements

This paper is the result of the analysis of a joint HMS team composed by CNES and ONERA. Putting together experts in propulsion and algorithmic domains, a continuous effort is being provided to propose new solutions and approaches to design and evaluate innovative diagnosis approaches for rocket systems.

References

[1] R. Isermann, "Process fault detection based on modeling and estimation methods: a survey", Automatica 20, 4 (1984), pp. 387-404.

[2] T. Boileau and N. Leboeuf and B. Nahid-Mobarakeh and F. Meibody-Tabar, "Online identification of PMSM parameters: Parameter identifiability and estimator comparative study", IEEE Transactions on Industry Applications, 47, 4 (2011), pp. 1944-1957.

[3] L. Liu and D. Cartes and others, "On-line identification and robust fault diagnosis for nonlinear PMSM drives", in American Control Conference (2005), pp. 2023-2027.

[4] C. Meyer and J. F. Zakrajsek, "Rocket engine failure detection using system identification techniques", in 26th AIAA/SAE/ASME/ASEE Joint Propulsion Conference, Orlando, FL (1990).

[5] Iannetti, S. Palerm, J. Marzat, H. Piet-Lahanier, G. Ordonneau, A new HMS for the Mascotte cryogenic test bench, AAAF, Space Propulsion Conference, Rome Italie, 2-6 May 2016

[6] Iannetti, J. Marzat, H. Piet-Lahanier, G. Ordonneau, Automatic tuning strategies for model-based diagnosis methods applied to a rocket engine demonstrator, PHM Europe 2016, Bilbao, Spain, 5-8 July 2016

[7] G. P. Sutton. History of liquid propellant rocket engines. AIAA, 2006.

DOI: 10.13009/EUCASS2017-417

[8] E. Chow and A. Willsky, "Analytical redundancy and the design of robust failure detection systems", IEEE Transactions on Automatic Control, 29, 7 (1984), pp. 603-614.
[9] M. Labarrère and J.P. Krief and B. Gimonet and M. Pélegrin and J.P. Krief, Le filtrage et ses applications (Cepadues, 1982).
[10] D.G. Luenberger, "Observing the state of a linear system", IEEE Transactions on Military Electronics, 8, 2 (1964), pp. 74-80.
[11] B. Larroque, "Observateurs de systèmes linéaires: application à la détection et localisation de fautes" (2008).
[12] D. Boutat, "La notion d'observateur pour les systèmes non-linéaires", Supméca pour la Journée Mathématique (JMS)
[13] M. S. Grewal, Kalman filtering (Springer, Berlin Heidelberg, 2011).
[14] A.Zolghadri, "Algorithm for real-time failure detection in Kalman filters", IEEE Transactions on Automatic Control, 41, 10 (1996), pp. 1537-1539.
[15] R. J. Patton, "Fault detection and diagnosis in aerospace systems using analytical redundancy", IET Computing & Control Engineering Journal, 2, 3 (1991), pp. 127-136.
[16] T. Kobayashi and D. Simon, Application of a bank of Kalman filters for aircraft engine fault diagnostics (Defense Technical Information Center, 2003).
[17] U. Lerneri and R. Parr and D. Koller and G. Biswas and others, "Bayesian fault detection and diagnosis in dynamic systems", in 17th National Conference on Artificial Intelligence AAAI, Austin, TX (2000), pp. 531-537
[18] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems", in AeroSense'97, Orlando, FL (1997), pp. 182-193.
[19] J. Crassidis and J. L. Junkins, Optimal estimation of dynamic systems (CRC press, 2011).
[20] M. NøRgaard, N. Poulsen, O. Ravn, "New developments in state estimation for nonlinear systems", Automatica 36, 11 (2000), pp. 1627-1638
[21] Z. Wang and X. Liu and Xiaohui, Y. Liu, J. Liang and V. Vinciotti, "An extended Kalman filtering approach to modeling nonlinear dynamic gene regulatory networks via short gene expression time series", IEEE Transactions on Computational Biology and Bioinformatics (TCBB), 6, 3 (2009), pp. 410-419.
[22] A.Falcoz, "Contribution au développement de stratégies de diagnostic à base de modèles pour les véhicules spatiaux : application à une mission de rentrée atmospherique", Ph.d Thesis (2009).
[23] K. Reif and R. Unbehauen, "The extended Kalman filter as an exponential observer for nonlinear systems", IEEE Transactions on Signal Processing, 47, 8 (1999), pp. 2324-2328.
[24] B. Sadeghi and B. Moshiri, "Second-order EKF and unscented Kalman filter UKF fusion for tracking maneuvering targets", in International Conference on Information Reuse and Integration, Las Vegas (2007), pp. 514-519.
[25] E. Wan and R. Van Der Merwel, "The unscented Kalman filter for nonlinear estimation", in Adaptive Systems for Signal Processing, Communications, and Control Symposium (2000), pp. 153-158.
[26] L. Lavigne and A. Zolghadri and P. Goupil and P. Simon, "Robust and early detection of oscillatory failure case for new generation Airbus aircraft", in Proc. of the AIAA Guidance, Navigation and Control Conference and Exhibit, Honolulu (2008).
[27] Ha, Q.P. and Negnevitsky, M. (1970). “Fuzzy tuning in motion control”. WIT Transactions on Information and Communication Technologies, 16.
[28] Guidorzi, R. and Marro, G. (1971). “On Wonham stabilizability condition in the synthesis of observers for unknown-input systems”. IEEE Transactions on Automatic Control, 16(5), pp. 499-500.
[29] Darouach, M., Zasadzinski, M., and Xu, S.J. (1994). “Full-order observers for linear systems with unknown inputs”. IEEE Transactions on Automatic Control, 39(3), pp. 606-609.
[30] Hou, M. and Muller, P. (1992), “Design of observers for linear systems with unknown inputs”. IEEE Transactions on automatic control, 37(6), pp. 871-875.
[31] Zhu, F. (2012). “State estimation and unknown input reconstruction via both reduced-order and high-order sliding mode observers”. Journal of Process Control, 22(1), pp. 296-302.
[32] Witczak, M. (2007). “Modelling and estimation strategies for fault diagnosis of non-linear systems: from analytical to soft computing approaches”, volume 354. Springer Science & Business Media
[33] Patton, R. and Chen, J. (1997). “Observer-based fault detection and isolation: Robustness and applications”. Control Engineering Practice, 5(5), pp. 671-682.
[34] Ryu, J.H.,Wan, H., and Kim, S. (2010). “Optimal design of a CUSUM chart for a mean shift of unknown size”. Journal of Quality Technology, 42(3), pp. 311-326.
[35] Sparks, R.S. (2000). “Cusum charts for signalling varying location shifts”. Journal of Quality Technology, 32(2), pp. 157-171.
[36] Le, K., Huang, Z., Moon, C.W., and Tzes, A. (1997). “Adaptive thresholding - a robust fault detection approach”. In Proceedings of the 36th IEEE Conference on Decision and Control, , volume 5, pp. 4490-4495.

[37] Emami-Naeini, A., Akhter, M.M., and Rock, S.M. (1988). “Effect of model uncertainty on failure detection: the threshold selector”. IEEE Transactions on Automatic Control, 33(12), pp. 1106-1115.
[38] Isaksson, A. (1992). “An on-line threshold selector for failure detection”. Linköping University Report.
[39] Basseville, M., Nikiforov, I.V., et al. (1993).” Detection of abrupt changes: theory and application”, Englewoods Cliffs: Prentice Hall.
[40] Mohanty, S., Pradhan, A., and Routray, A. (2008). “A cumulative sum-based fault detector for power system relaying application”. IEEE transactions on Power Delivery, 23(1), pp. 79-86.
[41] Page, E. (1954). “Continuous inspection schemes”. Biometrika, 41(1/2), pp. 100-115.
[42] Jiang, W., Shu, L., and Apley, D.W. (2008). “Adaptive CUSUM procedures with EWMA-based shift estimators”. IIE Transactions, 40(10), pp. 992-1003.
[43] Marzat J., Piet-Lahanier H., Damongeot F., Walter E. (2012) “Model-based fault diagnosis for aerospace systems”, Proc. Imech. Part G: Journal of Aerospace Engineering, 226(10), pp. 1329-1360.
[44] Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., and Yin., K. (2003). “A review of process fault detection and diagnosis Part III: process history based methods”. Computers and Chemical Engineering, 27(3), pp. 327–346.
[45] Hamayun, M.T., Edwards, C., and Alwi, H. (2013).” A fault tolerant control allocation scheme with output integral sliding modes”. Automatica, 49(6), pp. 1830-1837.
[46] J. Marzat, E. Walter, H. Piet-Lahanier, and F. Damongeot. (2010) “Automatic tuning via kriging-based optimization of methods for fault detection and isolation”. In Conference on Control and Fault-Tolerant Systems (SysTol), pp. 505_510.
[47] J. P. Herzog and R. L. Bickford Y. Y. Yue. (1998) “Dynamics sensor validation for reusable launch vehicle propulsion”. In 34th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Cleveland, OH, number AIAA-1998-3604.