Considerations for Obtaining Tangible Operational and Maintenance Benefits from Aircraft Health Monitoring Systems in a Big Data Environment

Normal Paper

Stephen R. Hall[1], Catherine Cheung[2] and John W.R. Miner[1]

[1] *Celeris Aerospace Canada, Ottawa, Canada;*
[2]*National Research Council of Canada, Aerospace Business Unit, Ottawa, Canada;*
halls@celeris.ca

Abstract

Rapid developments in Fly-by-Wire technology, data recorder processing and storage capability combined with technologies such as Hadoop, Big Data and advanced analytics have created a "data-tsunami" within the aerospace health monitoring industry. Organizations wishing to derive tangible operational and maintenance benefits from this data-tsunami need to address three main issues. First, they should implement effective data management and processing strategies. Second, they should ensure that each parameter identified for acquisition supports current and anticipated operational, maintenance and life-cycle management and fleet objectives. Finally, they should evaluate the actual cost-benefit of adding additional parameters and the associated impact on overall system design, implementation and cost. Effective solutions to these issues can only be obtained through a multi-disciplinary approach that evaluates and optimizes the relative merits and implications of proposed solutions from an overall system design perspective as opposed to an individual constituent "discipline-silo" perspective.

Keywords: Aircraft Health Monitoring, Big-Data, Predictive Maintenance, Proactive Maintenance, Advanced Analytics

Introduction

Rapid technological development on a number of fronts have created a "data-tsunami" within the aerospace industry. Thousands of parameters recorded during each flight can now be captured with the size of the resulting individual flight files ranging anywhere from 500 GB to 2-3 TB. Ongoing advances in technology coupled with the trend to "capture all available data" based on the assumption that one day it may prove valuable, suggests that the number of parameters captured (and the corresponding data file sizes) will continue to increase at almost exponential rates for the foreseeable future.

Unfortunately, the quantity of the data that is acquired and available for "analysis" has brought about significant challenges which can limit the tangible operational, maintenance and safety benefits that can be extracted from this data[1][2]. The primary challenge is that of not being overwhelmed with data where there is no clear understanding as to how it can be processed, validated, understood and interpreted. Subjective estimations obtained from informal discussions suggest that in many cases less than 5% of the data that is captured is being processed and used and that this may be a "high-end" estimate! While big-data and descriptive, predictive and prescriptive advanced analytics are often proposed to address this problem, their successful application is frequently handicapped by the lack of robust automated data validation tools capable of ensuring that the "clean-data", essential for meaningful and accurate predictions, are obtained[3].

A second, but not insignificant challenge arises from a widely-held and mistaken belief that the more data that is collected, the greater the insight can be obtained regardless of whether the rationale for collecting the data is understood [4][5]. The viability of implementing such an approach has been greatly aided by rapid advances and/or cost-reductions in Fly-By-Wire technology (simplifies data acquisition through the ability to rapidly connect to aircraft data-buses), advances in data acquisition technology (facilitating the capture and/or processing of thousands of parameters at high sampling rates) and the development of large, cheap data storage units.

The capability to acquire large numbers of parameters leads to a third challenge namely the belief that the acquisition of parameters is close to a "zero-cost" item as many of the current data acquisition systems can readily and economically be expanded and increased data-storage capacity of the acquisition systems is inexpensive. While these latter statements are true, they do not reflect the major data validation and storage costs associated with the process, storage and management of huge quantities of data. As discussed later in this document, from an overall program perspective, the processing, storage and management of data are definitely not zero-cost items. These costs are very significant and need to be factored into any business proposal that evaluates the anticipated Return-On-Investment (ROI) of any Aircraft Health Management Program (AHMP).

Effective solutions to these issues can only be generated through a multi-disciplinary approach that evaluates the relative merits and implications of proposed solutions from an overall system perspective as opposed to merely from the perspective of each of the constituent disciplines of which the system is comprised. An engineering solution that does not account for practical data processing constraints and a data processing solution that does not account for practical engineering functionality both result in the same outcome; a non-functional analytics capability which is unable to deal with the acquired data that is being transmitted to it!

This paper identifies some of the multi-disciplinary issues that need to be addressed to obtain tangible operational and maintenance benefits from an Aircraft Health Management Programs functioning in a "Big-Data" environment. A more in-depth discussion of the issues can be found in Reference [1].

Data Overload – A Tsunami in Waiting!

As illustrated in Table 1 the number of parameters that are captured from operational aircraft has grown, and continues to grow, almost exponentially.

Table 1: Growth in AHMS Data

Aircraft Model/Type	Introduced into Service	Number of Available Parameters	Comments
A-320	Mid 1980s	20,000[6][7]	
A-380	Mid 2000s	250,000	
A-350	End 2010s	400,000	
B 787	End 2000s	140,000[8]	Estimates of 1.4 TB/3 Hour Flight (includes, data, video, voice communications)
Latest Gen Regional Jets	Mid 2010s	100,000	Estimates of 300 MB to 1 TB per flight with average stage lengths around 1.5 – 2.0 hours. Typically, five flights/day.

The result is that analysts quickly become buried in data. For example, the latest generation of regional jets can accumulate up to 500 Gb of data during a 1.5-hour flight. Typically, these jets may undertake five flights of similar duration a day. Therefore, assuming an 80% weekly utilization factor to account for maintenance down-time etc., an airline operating a fleet of twenty of these aircraft could accumulate 280Tb of data per week! This is by no means a trivial amount of data! While many proponents argue that the management of such data quantities falls well within the scope of Big Data and distributed processing techniques such as Hadoop, the management of the data is only one part of the problem. The work required to evaluate how each piece of data is to be processed, validated, interpreted and its implications disseminated so that timely and preventative operational and/or maintenance actions can be implemented is also far from trivial[1]. Of particular concern is data validation. The quantities of data that are being collected are so huge that many organizations seem to be assuming that if the data is formatted correctly, it is also validated. Nothing could be further from the truth[3]! Careful consideration of what constitutes valid data, how erroneous data will be detected and the timely actions needed to rectify invalid data and correct any associated sensor/hardware problems is essential. Generation of such criteria is time-consuming and best addressed through the generation of Use Cases[1][9]. Such use cases will not only address simple single parameter range checks but also need to address multi-parameter consistency checks[1].

Data Insight – There is no Magic Bullet!

Aircraft parameter acquisition first started in the late 1950's. Initially AHMP scope was primarily limited by the available data acquisition and storage technology; a situation that persisted into the early 1980's[10]. Due to the limited number of parameters that could be reliably acquired, it was imperative that careful consideration be given to the reason(s) for acquiring each parameter, the anticipated results and how well the acquired data would accurately characterize the response of the structural or system components of interest. Rapid technological advances on several fronts over the last twenty years have essentially removed these technology limitations. Additionally, the acquisition costs of systems and sensors capable of monitoring thousands of parameters via aircraft data buses has dropped considerably. Finally, concurrent developments in data analytics ostensibly provide powerful tools which it is argued removes much of the need to consider the rationale for acquiring each parameter as the underlying relationships can be discovered through the application of advanced statistical and data-analytics techniques. This appears to be leading to an increasingly pervasive and mistaken belief that all available parameters should be captured, regardless as to whether a reasonable expectation as to how they might provide insight is apparent, because "more data is always better, as it might prove to be useful at a later date". An implicit assumption arising from such a misunderstanding is that because powerful statistical and data analytics tools capable of mining "anticipated and unanticipated (buried)" correlations are available, the data itself will reveal "meaningful" relationships upon which further investigation should be focussed. Ultimately, this ignores that while correlation is a necessary condition to establish causality (relationship) between parameters and aircraft utilization of interest, it is not a sufficient condition (i.e.: because correlation is established it does not mean that a significant relationship or problem has been identified)[11].

While modern-day statistical and data analytics tools provide powerful capabilities that can be used to obtain significant insight into AHMS parameter relationships, they do not replace the need for the careful and systematic consideration and understanding of the objectives and design of AHMS programs by domain experts (engineers). Additionally, if tangible operational, maintenance and life-cycle management Return-on-Investment(s) (ROI) are to be obtained, it should be readily apparent how each parameter identified for acquisition supports corporate (Owner/Operator) current and anticipated operational, maintenance and life-cycle management and fleet objectives[1][5][10]. Collecting all parameters available without an understanding of their actual or anticipated benefit will significantly increase the risk that little-to-no tangible ROI will be obtained from these programs[1][12]. Instead of acquiring all parameters "just in case they are needed to fulfil some ill-defined objective(s) downstream", careful consideration should be given to selecting meaningful, or at least potentially meaningful, parameters that will result in useful or wise data being obtained where wise data is defined as[13]:

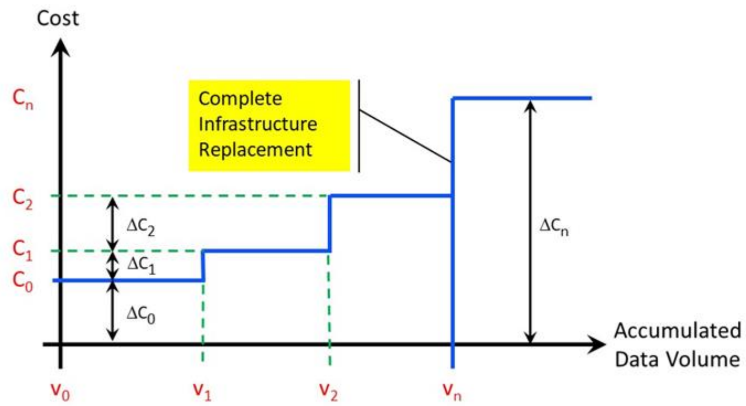
> *"Data that based on physical principles has been thoughtfully identified as having the potential to shed light on a given problem and/or identify potential causal factors that may result in structural or system failures."*

More detailed discussions pertaining to the definition and selection of "Wise-Data" can be found in References [2] and [13].

Data Acquisition – Understanding the Real Costs!

As alluded to earlier in this paper, while the apparent costs associated with parameter acquisition appears to be relatively small, over the life of a program, it is not the parameter acquisition that is the main cost. The data volumes that are now being received are so huge that these volumes are taxing existing data management infrastructure and requiring significant investment in upgraded infrastructure. If the current rate of growth remains unabated, it will not be long before storage requirements will be sized in Exabytes (EB - 10^{18}). Aside from the cost of data transmission, significant infrastructure costs will also be incurred for support and management of this volume of data. This is illustrated schematically in Figure 1.

Figure 1: Data Volume Versus Infrastructure Cost



When setting up any AHMS data management program, there will be an initial cost that has to be incurred (ΔC_0). Typically, some growth (system scalability) capability is included in the initial infrastructure. Consequently, as the data volume grows, the initial infrastructure can cope with this growth through modest additional investment in additional processing capability, storage capacity etc. (ΔC_1). Such modifications may occur several times (ΔC_2). However, at some point a combination of technology obsolescence and ever-increasing data volumes will require the acquisition of a completely new infrastructure and, in many instances, the need to migrate data from the original infrastructure to the new infrastructure. The additional investment required (ΔC_n) can be quite significant[1].

If the data that is driving up the infrastructure costs is relevant and will generate tangible operational, maintenance and life-cycle management ROI which yields an overall positive cost-benefit, then such infrastructure expenditure is appropriate. However, if the vast majority of the data is merely being acquired because it might "someday be useful", any potential ROI may easily be negated through the incurrence of ongoing and unnecessary infrastructure costs[1].

Data Utility – Adopt an Integrated System Approach!

AHMP programs are multi-disciplinary in nature and may require input from Subject Matter Experts (SMEs) in such disciplines as aircraft structures, aircraft systems, signal processing, remote/real-time data acquisition, big-data, block-chain, data analytics, aircraft operations and aircraft maintenance[1]. Additionally, any AHMS system is liable to be comprised of several different sensors and systems from different vendors and be required to integrate with a number of existing aircraft management systems such as aircraft configuration databases, reliability-centered maintenance databases, electronic aircraft log-books etc., all of which may or may not be incorporated into an Enterprise Resource Planning system[1]. Consequently, the successful implementation of an AHMP that produces tangible operational, maintenance ROI requires a disciplined approach to overall system integration[1][10]. While essential, this can prove to be challenging in a world where the complexities of many of the constituent technologies make it tempting to conduct system development in isolated "discipline design silos" as organizations seek to "focus on their "core expertise".

Concluding Remarks

Ongoing advances in technology coupled with the trend to "capture all available data" based on the assumption that one day it may prove valuable, suggests that the number of parameters captured (and the corresponding data file sizes) will continue to increase at almost exponential rates for the foreseeable future. Organizations wishing to derive tangible operational and maintenance benefits from this data-tsunami need to address three main issues. First, they should implement effective data management and processing strategies to avoid becoming overwhelmed with data which they cannot process, validate, understand and interpret. Second, they should counter the widely held belief that the more data that is collected, the greater the insight that can be obtained regardless of whether the rationale for collecting the data is understood or not. Finally, they should evaluate the actual cost-benefit of adding additional parameters and the associated impact on overall system design, implementation and cost. Effective solutions to these issues can only be obtained through a multi-disciplinary approach that evaluates and optimizes the relative merits and implications of proposed solutions from an overall system design perspective as opposed to an individual constituent "discipline-silo" perspective. Consequently, the successful implementation of an AHMP that produces tangible operational, maintenance ROI requires a disciplined approach to overall system integration.

References

1. Hall, S. R. and Miner J.W.R., "Considerations Related to the Application of Predictive Analytics in an Aerospace Health Monitoring Environment.", *Celeris Aerospace Canada Inc. Technical Report, CAC/TR/18-002,* July 2018.

2. Hall, S. R., "We Don't Need Big Data; We Need Wise Data", *LinkedIn Article available from:* *https://www.linkedin.com/pulse/we-dont-need-big-data-wise-stephen-hall/*, October 27, 2015.

3 Hall, S.R., "Engineering Big Data Validation – Neglect at Your Peril!", *LinkedIn Article available from* *https://www.linkedin.com/pulse/engineering-big-data-validation-neglect-your-peril-stephen-hall/*, August, 19 2017.

4. Provost, F. and Fawcett, T., "Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking – 1st Edition", *Kindle E-Book. Available from: https://www.amazon.ca/Data-Science-Business-Data-Analytic-Thinking-ebook/dp/B00E6EQ3X4/ref=sr_1_1?s=digital-text&ie=UTF8&qid=1524063238&sr=1-1&keywords=data+science+for+business.*

5. "HBR Guide to Data Analytics basics for Managers", *Harvard Business Review Press, (ISBN 9781633694286),* 2018.

6. "Airbus Will Monitor all A 35 XWB Delivered to Customers Around the Clock", *A350 XWB News article, 20 August 2014. Available from http://bloga350.blogspot.ca/2013/08/airbus-will-monitor-all-a350-xwb.html.*

7. "Big Data Capabilities Boost Manufacturer Aftermarket Ambitions", *Flight Global, 6th December 2017. Available from: https://www.flightglobal.com/news/articles/big-data-capabilities-boost-manufacturer-aftermar-443055/.*

8. Humphries, M., "The Boeing 787 Produces Over 500 GB During Every Flight", *Geek.com Article, March 2013. Available from: https://www.geek.com/geek-cetera/the-boeing-787-produces-over-500gb-of-data-during-every-flight-1542105/.*

9 Hall, S.R., "Implementing Advanced Analytics in Legacy Aircraft Fleets", Celeris Aerospace Canada Inc. Presentation, 2018. Available From: *http://bit.ly/AHMSFDRLEGPDF*.

10. Hall, S. R., "Moving Beyond Aircraft Health Monitoring Mythology: The Journey Toward Obtaining Tangible Life-Cycle Management ROI". *Presented at the Sixteenth Australian International Aerospace Congress (AIAC 16), HUMS Conference, 2015. Available from: http://www.humsconference.com.au/Papers2015/Peer_reviewed/Hall01.pdf*.

11. Vigen, T., "Spurious Correlations", *Hachette Books, New York/Boston, May 2015. Available from: https://www.amazon.ca/Spurious-Correlations-Tyler-Vigen/dp/0316339431/ref=sr_1_1?ie=UTF8&qid=1524167790&sr=8-1&keywords=spurious+correlations.*

12. Hall, S.R., "Analysis of the Data Obtained from a Survey on Aircraft Structural Health Management/Monitoring Programs", *Celeris Aerospace Canada Inc. Technical Report CAC/TR/14-001, 16 June 2014. Available from: https://celerisaerospace.box.com/s/rpnwh0u6t6oz90brz5ta*

13. Hall, S.R., "Guidelines for the Identification and Management of Wise Data", *Available from: https://www.linkedin.com/pulse/guidelines-identification-utilization-management-wise-stephen-hall/* , January 3, 2018.