Change Point Analysis applied to SMS

By **Paulo Manoel RAZABONI**, EMBRAER Air Safety Department

RAZABONI *is graduated in Electronic Engineering, with a specialization degree in Air Safety, Production Management and Administration. He is an Accredited Element for Aeronautical Accidents Prevention and Investigation System (SIPAER – Brazil). After joining EMBRAER in 2007 at Technical Support area, he moved to the Air Safety Department in 2010, where he is nowadays the Data Analysis team manager. This team is responsible for analyzing flight data, either on a regular basis, or providing support to investigations. Other roles include generating safety-related statistics for the fleet and helping in the development of specifications for future aircraft data recorders.*

Abstract

Dealing with on-site accident investigation will always be a significant part of every safety system. Nevertheless, as aircraft systems become even more precise and reliable, thus pushing safety performance indexes to higher standards, the need for ways to improve event precursor's detection becomes paramount in order to keep absolute numbers within an acceptable limit despite the ever increasing volume of flight operations.

As no one could wait for an accident to occur to start fighting its contributors, a proven way to cut these rates is to work in the "base of the pyramid", exploring incidents, reports or even routine operations in search for the so-called "lower-consequence indicators".

The present challenge is how to share the always limited efforts among all probable issues and, especially inside the Safety Management System (SMS), how to generate reliable safety performance indexes related to the "never-happened" events. While the experience gathered in the field is very relevant to the prevention job, the monitoring of selectively identified items is essential to assure adequate risk mitigation, which in turn collaborates in achieving an acceptable level of overall safety performance.

An objective tool to monitor selected events rates will be shown, using statistical ways to positively identify changes in their behavior, specially focused on "low occurrence" events. Detected changes may be assigned to the negative effect of some factor getting out of control, or to the positive effect of an applied countermeasure. In the first case, an alert can be issued and, in the second case, the effectiveness of the actions can be better understood and measured, and a very accurate index can be assigned to a safety performance key indicator.

The technique fundamentals will be explained, illustrated with some cases, and a hands-on spreadsheet will be shared.

Introduction

EMBRAER, as other aircraft manufacturers, gathers information about events related to Air Safety from operators' reports and other sources, as government authorities and specialized companies. These reports are grouped as per their nature, for example ATA chapter or FAA Nature of Condition codes.

Monthly, or as per some special requirement, the information is summarized and sent to a distribution list, which includes mainly people from Customer Support, Systems Engineering, and an Air Safety team responsible for the "Safety Health" monitoring. These summaries include, besides the data grouped in Pareto and Pie charts, the history and a statistical trend analysis, developed specially for this kind of data, as will be explained herein. This analysis eliminates the subjective interpretation of data, supporting resource allocation for crucial tasks, as well as confirming the effectiveness of some corrective action taken, so closing the loop for a change initiative.

The techniques were chosen among several sources as per the specific data nature, as a way to gather the information that would better fit the decision process (HITOSHI, 1993). The main tool is a "Change Point Analysis" approach, as described in one of the references (TAYLOR, 2012). This analysis allows estimating the precise point where the "process" represented by the input data changed its behavior, thus calculating representative figures for the average and deviations for each segment (before and after the change itself).

Each confirmed change point found brings up another question, that is if there may be another "minor" change point, that would become relevant after isolating the primary one. So, each segment can be recursively checked for the presence of a secondary point that fulfills the same requirements. During the analysis validation, it was assumed that only the change points with at least 95% of confidence level would be considered, with the method for calculating this index to be explained also.

As a way to summarize all this information, a process-like chart was chosen (STAPENHURST, 2005), using symbols and colors to make it friendly enough for all the recipients.

Average and deviation lines are drawn for each data segment, thus evidencing the changes detected, as well as short- and long-term trend lines. A distribution profile helps supporting the process chart analysis. Seasonality can be assessed by calculating the monthly stratified average for the last three years, then interpolating a sinusoidal curve that would better fit to data.

Methodology

For the change points detection, an algorithm called "Cumulative Sum of Differences" to the average was used (OAKLAND, 2008). Then, for a particular data set, one could get the following history chart:

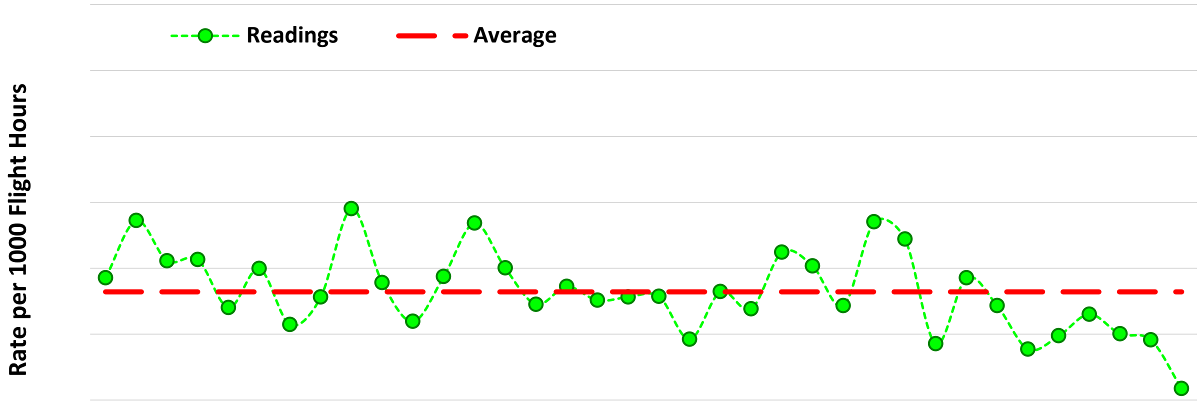

***Figure 1**. Typical data input, as a rate against time (in months).*

Although seeming to be simple, determining a change point from the above data can be very challenging. Several trend analyses may be tried. For instance, some points sequentially above or below the average. Although all techniques could be implemented by software, some of them would result in different outcomes, and a numeric value for the change itself would be hard to get... Calculating the Cumulative Sum would produce the output:

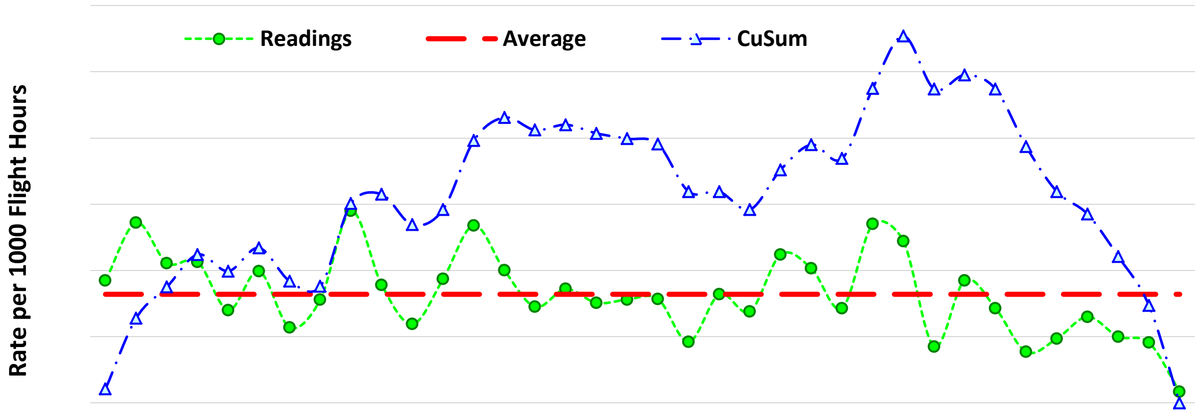

***Figure 2**. Cumulative Sum of differences to the average for the data set.*

The chart begins with the difference between the first reading and the average and necessarily ends at zero, as by definition the sum of all differences to the average is zero. The x-coordinate where the value is the farthest form the average defines a candidate for change point. Conceptually, from this point on, readings shall contribute in a different way to the average.

At this moment, it becomes necessary to check whether the original readings present a distribution pattern significant enough, not only a mere sequence of random values that might produce a peak somewhere. The most straightforward way to do so is to force this situation, shuffling the data and checking if some random distribution would be able to generate comparable results (in this case, a higher amplitude in Cumulative Sum curve). As per one of the references (TAYLOR, 2012), proceeding this way a thousand times shall be a fair enough to classify the candidate as a change point or not. For example, if data is randomly shuffled and cumulative sum curve amplitude is calculated 1,000 times, and for 950 times the amplitudes remain below the original one, the candidate can be considered a change point with 95% of confidence. In an analog way, if no other distribution was able to produce higher amplitudes, the candidate would be considered a change point with 100% of confidence. In this example, a change point could be assigned to the process, as follows:

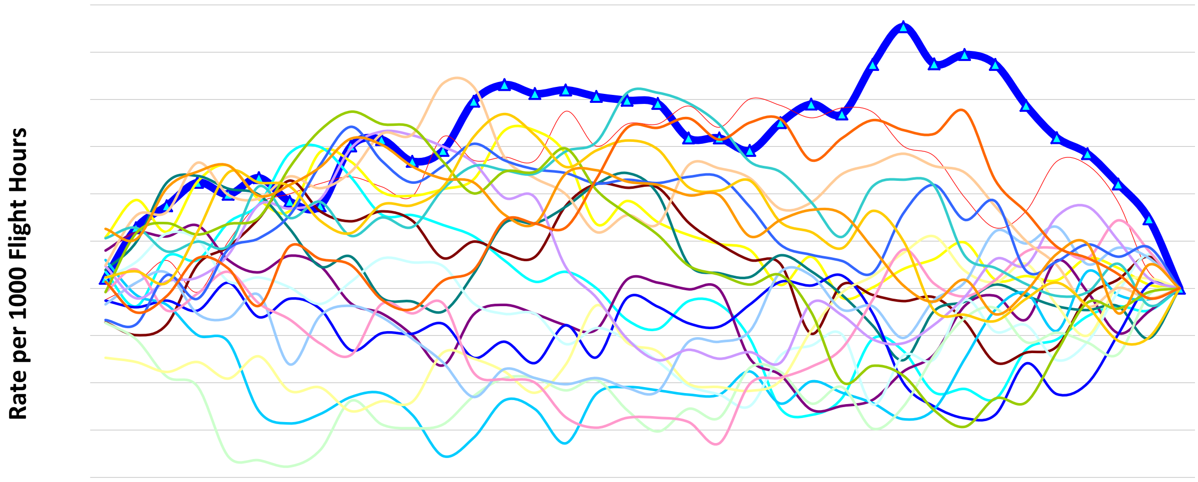

Figure 3. *Cumulative Sum of differences to the average, shown here for the original data set (bold blue line with markers), and for twenty other calculations. Amplitude of each curve is compared to the original one, for confidence level determination.*

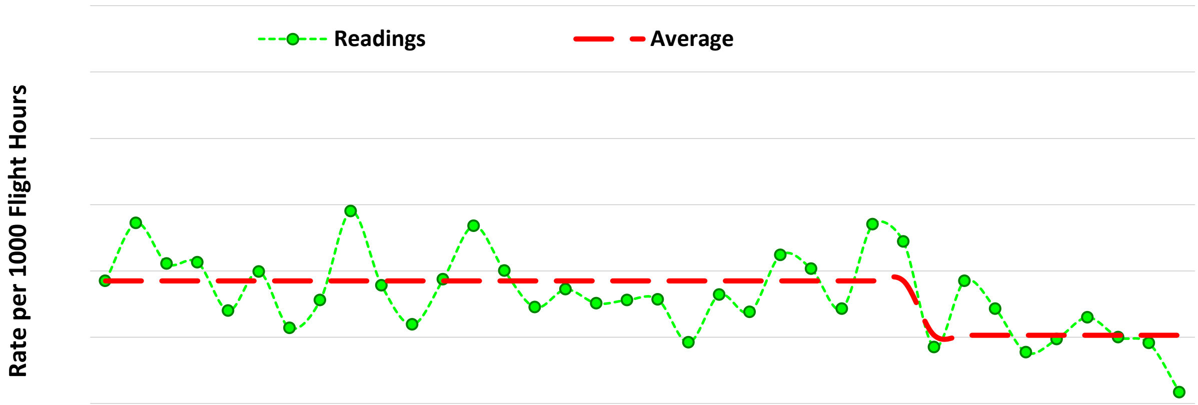

Figure 4. *As the candidate was positively classified as a change point, average calculation is*

performed for both data segments (before / after it). Then, the same proceeding is applied to each segment, recursively.

Assuming that data revealed a behavior change, a similar analysis can be run again twice, as now we have two distinct data segments, one before and other after the change point. Applying the same technique recursively will reveal other change points, if any, up to a practical limit, which may be one of the two following cases: no more candidates could be found with at least 95% of confidence; or there are no enough points left in the segment to drive a significant analysis.

From the knowledge about change points for the parameter under study, one can better decide about resources allocation, as well as check the effectiveness of some action taken.

This technique is resilient enough against spurious readings (TAYLOR, 2012). In order to help evidencing the behavior, a process chart, called “u-chart” was chosen, which allows variable opportunity for the events to happen (as Flight Hours usually vary from month to month).

The deviation "s" can be calculated as below (STAPENHURST, 2005):

$$s = \sqrt{\bar{u}/n} \qquad \text{(Equation 1)}$$

$\bar{u}$ is the average for the data segment (as identified from the change point analysis);

n are the Flight Hours for that specific month

As “n” may vary, 2-s and 3-s limit curves are not purely horizontal, but they show amplitude variation that is inversely proportional to the hours flew during each month. These two limits are usually called “warning threshold” (2-s) and “action threshold” (3-s) for processes under control.

In order to provide a consistent view over the “process chart” (as it is valid only for “normal-like” distributions), it is necessary to know the readings’ frequency distribution, which may be done by plotting a histogram, in this case drawn 90 degrees rotated and placed aside the control graph using the same vertical scaling, a very common practice for simultaneous viewing (STAPENHURST, 2005).

Now, in order to evidence seasonality, a sinusoidal interpolation with a fixed 1-year period can be plotted over monthly-stratified average readings. This means that three Januarys, three Februarys and so on are averaged representing one typical month, as a way to get rid of long-term trends. A correlation coefficient is calculated, as well as the sinusoidal interpolation amplitude referred to the global average, which together will serve as a measure for the seasonality fitness and dependence for each parameter, which would help someone to take countermeasures in advance, as a bonus.

Putting all information together may seem to be challenging. So, the correct choice for colors and symbols is evident. Of course, as it is usually not required for all the recipients to deal with such level of details, a brief comment in plain text from the analyst for each plot would be appreciated. The curves will remain as an analysis tool for the specialists, and a source for deeper information, if required.

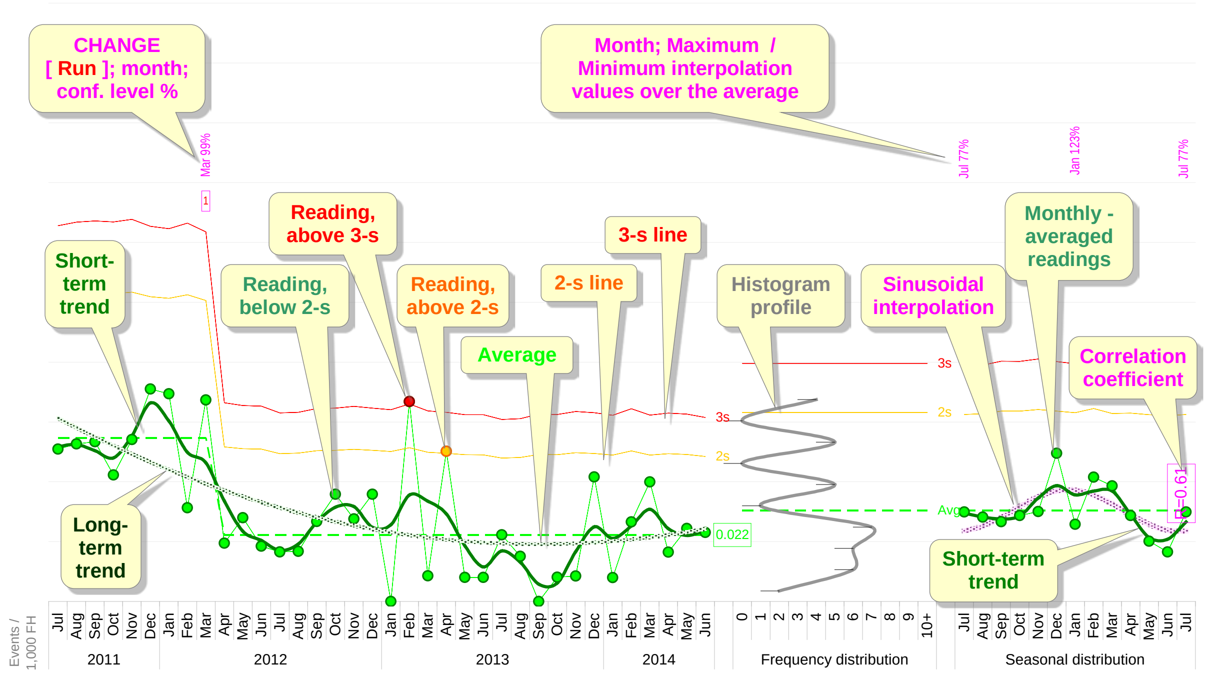

Figure 5. *Proposed chart, containing all elements with a brief description for each one. Styles were chosen to be intuitive, using standard representation when practical.*

The main throughput is that calculations are made automatically and the analyst expertise is focused on data interpretation supported by statistical tools, assertively documenting the process already done and generating targets for the next improvement cycle. Actions can be requested inside the company or from suppliers, in an unambiguous way. In the case of SMS implementation, performance indexes can be tracked to achieve established levels.

While the results for the manufacturer are clear, and the documentation would be concise enough to attend the Safety Management System requirements, the final client would also have benefits, as Safety as a whole would be closely monitored and improved.

The same technique may be applied also for monitoring items not directly related to Safety, as the ones related to maintenance, which can give expressive payback, like engines, landing gears, avionics subsystems etc.

All processing can conveniently be implemented by software, using dedicated routines in a database system, or even an ordinary spreadsheet for proof of concept evaluation, where Macro routines can be programmed (Macros are necessary, mainly because of the intense and recursive calculation, but will take only a few seconds to run). An already implemented template sheet is available, serving as a base for future developments or improvements, and can be requested to the author.

Examples

As it is impossible to control something that is not known, looking at the behavior of selected monitored items over time can give a rough idea about whether an intervention is necessary, as well as whether some action was successful.

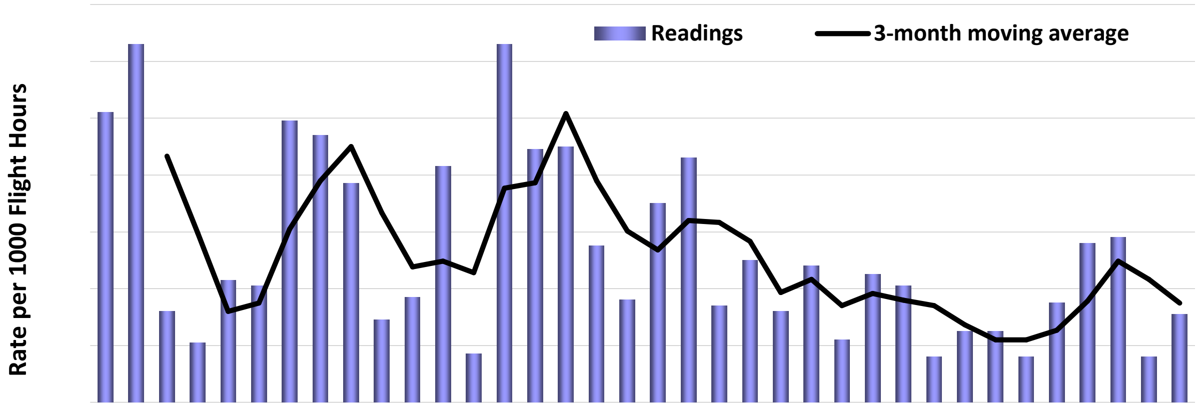

Figure 6. *Real data plotting, along with a short-term trend line based on the rolling average for the last three readings. Although some visual clues may suggest trends, subjectivity must be taken away in order to get an accurate analysis.*

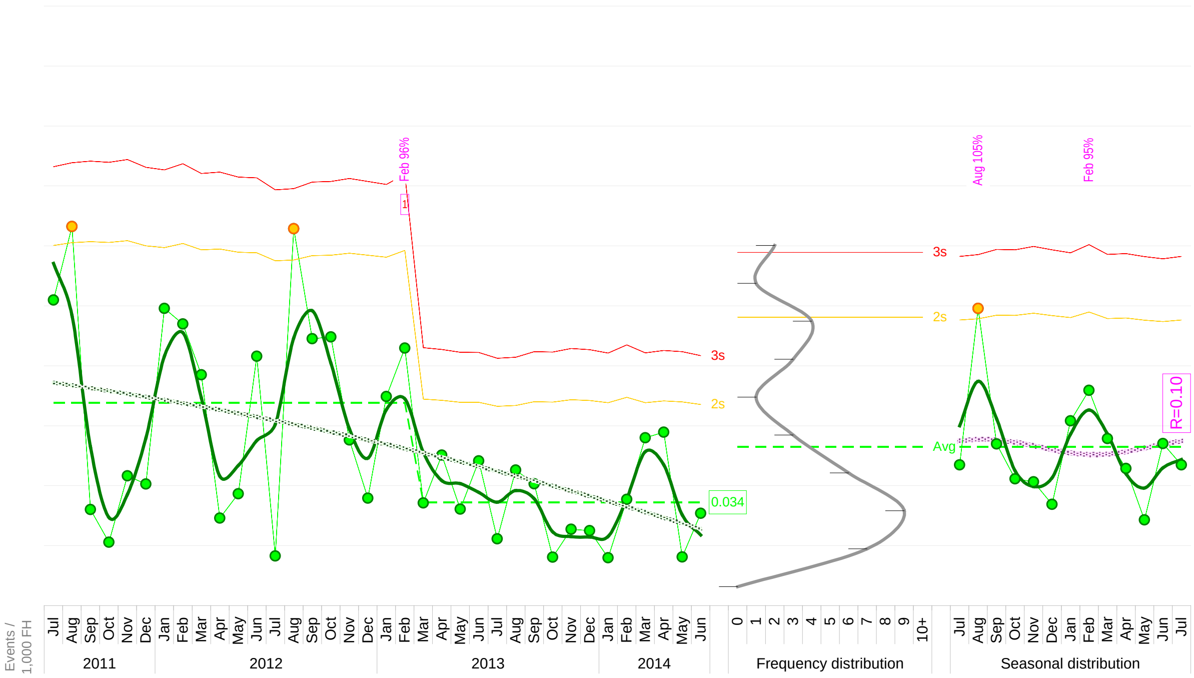

Figure 7. *Same data set, plotted using change detection algorithm. One "level 1" change was detected in Feb/2013 (with 96% confidence). Short-term behavior has become more stable (less scattering), while long-term curve presents a downtrend. Distribution is not normal for the whole set, but the peak is under the average, probably due to the recent readings. No seasonality was evidenced, as correlation factor is very poor.*

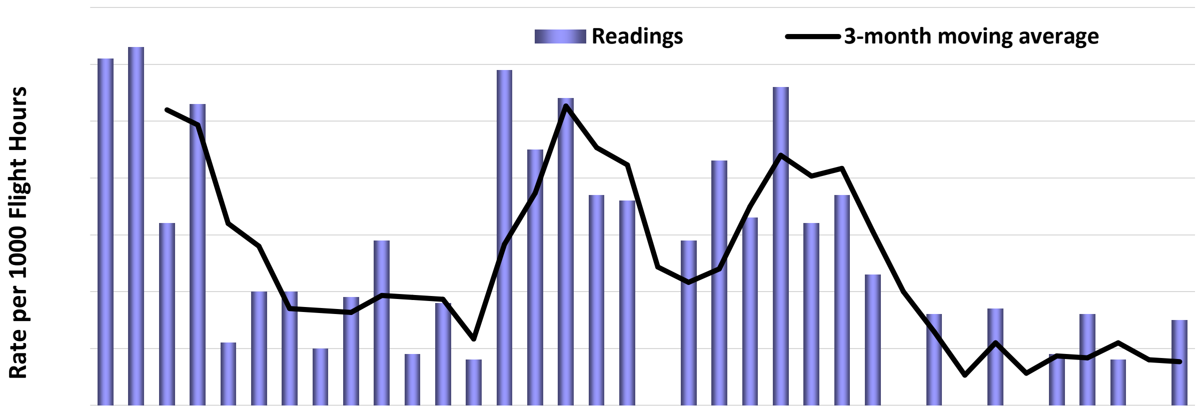

Figure 8. *Again, real data plotting. How many changes would be assigned to this data set?*

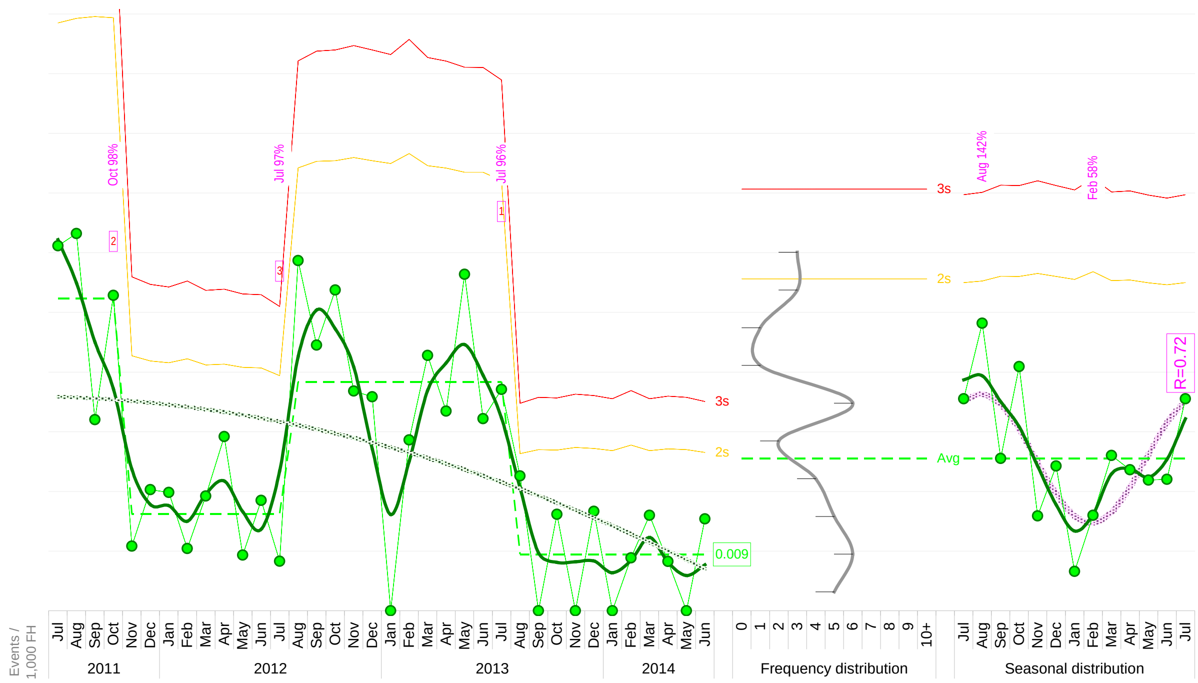

Figure 9. *Using the above data set, the algorithm reveals three changes, the latest one in Jul/2013 (96% confidence), earlier ones in Oct/2011 and in Jul/2012. Short-term behavior is stable, and long-term presents a downtrend. Distribution is not normal considering the whole set. Some level of seasonality may be assigned, as the sinusoidal curve roughly fits the data, the worst month identified as August (42% above the average).*

People with knowledge on the subsystem would be able to illustrate this analysis with elements from product history. As an example, average changes might suggest deviations or improvements, while seasonality might be associated with climate changes. The decision for assigning engineering resources, or sending communications (Service Bulletins, Newsletters etc.) to the operators would be facilitated and supported. This would also happen to the negotiations with suppliers, as a way for enforcing corrective actions, writing dispositions or even evaluating contract clauses.

Conclusion

This paper offers a method to better understand the behavior of safety-related items, using a statistical approach, in order to support strategic decisions like resources assignment or product change requests. A spreadsheet with Macro routines for fast data processing is available from the author. It was created to fit specific data characteristics, which are a common sense for commercial aircraft operations, and has its own benefits and limitations. Although it is a belief that this is a valuable tool, the suitability for solving specific tasks, related or not to Safety, must be adequately evaluated by the analyst.

Anyway, statistical methods, if adequately implemented, will provide reliable indicators which can be used to keep processes under control and, of course, to document the relevant actions taken in any quality system, including SMS.

This work is part of EMBRAER permanent commitment to the product safety in all levels, from concept to operation. As always, comments and suggestions are welcome, as a way for the process continuous improvement.

Acknowledgements

I wish to acknowledge the opportunity provided by ISASI for sharing this information, believing this can offer means to help improving aviation safety levels worldwide. My special thanks are extended to the staff of EMBRAER for the valuable feedback on the reports that my team has provided.

References

Kume, Hitoshi, **"Métodos Estatísticos para a Melhoria da Qualidade",** São Paulo - Brasil, Editora Gente, 1993

Oakland, John S. **"Statistical Process Control"**, Burlington/MA - USA, Elsevier Butterworth-Heinemann, 2008

Stapenhurst, Tim **"Mastering Statistical Process Control - A Handbook for Performance Improvement Using Cases"**, Burlington/MA - USA, Elsevier Butterworth-Heinemann, 2005

Taylor, Wayne A., **"Change-Point Analysis: A Powerful New Tool for Detecting Changes"**, in <http://www.variation.com/cpa/tech/changepoint.html>, accessed on August 12th 2014.