



Data Mining to Drastically Improve Spacecraft Telemetry Checking: An Engineer's Approach

David Evans¹, Jose Martinez² and Moritz Korte-Stapff³
ESA/ESOC, Darmstadt, Germany, D-64293

Attilio Brighenti⁴ Chiara Brighenti⁵ Jacopo Biancat⁶
S.A.T.E. Systems and Advanced Technologies Engineering S.r.l., Venezia, 30135, Italy

The number of telemetry parameters in a typical spacecraft is constantly increasing. At the same time the number of operators allocated to each spacecraft to check those parameters is constantly decreasing. Techniques such as limit checking are well known but they take time and effort to define, enter and manage as the mission evolves. The result is that the vast majority of telemetry parameters are not limit checked in real-time. In 2014, the Advanced Operation Concepts Office at ESA/ESOC decided to see if we could change this by employing Big Data type techniques on the data. The idea was simple, we asked our partner, SATE of Italy, to define future checks for all telemetry parameters given one year's worth of historical data. No engineering knowledge was provided and the derivation of the checks had to be completely automatic i.e. the checks had to be derived solely on the data itself with no human intervention. The mission we choose was Venus express (VEX) and the learning period ended just before the aero-braking activities started. We then applied these checks to the following three months of data which included interesting activities such as aero-braking preparation and aero-braking itself. This test data was not provided to SATE until after they had submitted their checks to us for validation. This paper describes SATE's response to this challenge. SATE decided to take a very pragmatic, engineering view of the problem and defined algorithms to search for anything that could be classed as constant in the data. This could be simple features of the data such as average or more exotic features such as harmonic mean, FFT coefficients and features

¹ Project Manager, Special Projects Division ESA/ESOC, david.evans@esa.int

² Data Scientist, Special Projects Division ESA/ESOC, david.evans@esa.int

³ Project Assistant, Special Projects Division ESA/ESOC, david.evans@esa.int

⁴ President, S.A.T.E. S. Croce 664/A Venezia Italy, attilio.brighenti@sate-italy.com

⁵ Group Leader, R&D - Diagnostic systems, S.A.T.E. S. Croce 664/A Venezia Italy, chiara.brighenti@sate-italy.com

⁶ Systems Identification and Control Engineer, R&D - Diagnostic systems, S.A.T.E. S. Croce 664/A Venezia Italy, jacopo.biancat@sate-italy.com

characterizing the sampling rate. 47 features were selected in the end (35 for numerical and 12 for categorical parameters), over different time windows resulting in over 500,000 possible time series. SATE delivered checks for every telemetry parameter of the VEX satellite, extending also the study to telemetry data of the XMM satellite available from a preceding project. This paper then goes on to describe the validation exercise carried out at ESOC in which the delivered checks were run on the new data and the results compared to actual operational events. After some optimisation, which were required to reduce the level of false negatives to reasonable levels the validation team produced some extremely interesting results creating a very accurate and detailed insight into the future operations. ESOC is currently planning to deploy these techniques operationally for flying spacecraft in the near future.

Nomenclature

<i>ESA</i>	=	European Space Agency
<i>FCE</i>	=	Flight Control Engineer
<i>FETCH</i>	=	Feature Checkability
<i>MUST</i>	=	Mission Utility & Support Tools
<i>SATE</i>	=	Systems and Advanced Technologies Engineering S.r.l. (private company located in Italy)
<i>SCOS</i>	=	Spacecraft Operating System
<i>SV</i>	=	Symptomatic Variable
<i>TWD</i>	=	Time Window Duration
<i>VEX</i>	=	Venus Express Satellite
<i>XMM</i>	=	XMM-Newton satellite

I. Introduction

The time honored method that operations engineers use to check that the mission (space and ground control) is performing as expected is by implementing checks in the operational databases. These are typically checks to see if spacecraft parameter values are between predefined limits or in predefined states. There are also command verification checks contained in the procedures or integrated into the control system. These are checks to see if spacecraft parameter values are between predefined limits or in predefined states after a command has been sent. Two important trends to note are that the number of spacecraft telemetry parameters and telecommands keeps increasing with time and the manpower for operating the ground systems keeps being reduced. This means that the present systems, involving engineers defining these checks individually and manually in operational databases/procedures, are by nature incomplete. What is needed in the future is a system that can derive checks automatically from historical data. Once such an automatic system exists it can be expanded to check many more spacecraft telemetry parameters, verify more commands and check more procedure steps with very little effort on the ground side. It can also be used to check much more useful features in the telemetry than the present limit and status checks employed.

In 2011 ESA took an important step in this direction by deploying a sophisticated checking algorithm called Novelty Detection. This can check combinations of some statistical features of some telemetry parameters against past history and detect if anything has significantly changed. This might not be an anomaly because it could be expected due to a change of mode or orbit etc. so the term “novel” was coined to mean “significantly different than the past”. The statistical features used as input for these numerical parameters were minimum, maximum, average and standard deviation and ESA has had some considerable success in using this operationally. In 2014 ESA started two parallel studies which were to build on this idea. Several questions were asked of Novelty.

- 1) Why only choose minimum, maximum, average and standard deviation as input features? There are many other ways to describe the features of a time series e.g. noise, rate of change etc. It may be that parameters will have some features that are useful to check, others that are not and this will be different for each parameter. This leads to the concept of a useful parameter-feature pair and associated check.
- 2) Why should we expect these checks to be always valid or always the same? Some parameter-feature pairs are expected to be only useful under certain conditions e.g. in eclipse, when a unit is in a specific state, when the spacecraft is communicating with the Earth, when the solar aspect angle is over a specific value on a particular panel.

- 3) Why only restrict the checks to numerical parameters? The majority of the parameters are status. All parameters should be included.
- 4) Why only check spacecraft telemetry in isolation? There may be strong relationships between other mission data sets and the behavior of the spacecraft data e.g. command files, mission planning data, or
- 5) Why only use this for routine spacecraft telemetry checking? Since telemetry checking is an integral part of procedure writing and command verification checks this should also be included.

II. The Objectives

The main objective of the study was therefore to analyze historical mission data (this included spacecraft housekeeping data but also ground events, orbit and attitude files, command history) and to search for new features and checks that would be useful in the telemetry checking, command verification and the procedure writing processes.

This was broken down into the following:

- To take a fresh look at the stored housekeeping data for a selected ESA mission and using the given data set to identify new types of features for each individual parameter that it would be useful to check in the future. Define the associated useful check to be applied for each parameter-feature pair.
- Using the given data set to identify parameter-feature pairs that change behavior over time taking on a limited set of values i.e. modes. Define the associated useful checks to be applied for each parameter-feature-mode pair.
- Look at mission data sets besides spacecraft telemetry (orbit determination files, command history etc.). Derive parameters whose evolution may have an impact on spacecraft telemetry parameter behavior. Process those mission data sets to create new time series data that detail the time evolution of these derived parameters in the same format as the spacecraft telemetry data.
- Identify linear cross-correlations among this global mission-related time series data (telemetry, derived and features) that are useful and define the associated checks.
- At each stage identify the necessary strategies to deal with the massive data sets involved.

III. The approach

The question was how to achieve and test these broad objectives? We came up with a simple idea. We asked the study partner, SATE of Italy, to analyze one year of historical mission data (telemetry, TC history, orbital files, SCOS logs, event packets) and to define useful telemetry checks that should be applicable in the following period. However there were rules.

- 1) No engineering knowledge was provided and SATE were not allowed to use any engineering knowledge (even derived) when defining the check. This was to ensure that whatever process was used would be immediately applicable to other missions without modifications being needed.
- 2) The derivation of the checks had to be completely automatic i.e. the checks had to be derived solely on the provided data with no human intervention.
- 3) No access to the data in the following period would be provided – SATE would only receive ESA’s assessment of how their checks performed during the study.

The mission we choose to provide the data was Venus Express (VEX). We choose a period of a year that ended just before the aero-braking activities started in 2014. Our plan was that once the partners submitted their checks we would apply them to the following three months of data which included novel activities such as aero-braking preparation and aero-braking itself. If the checks were really useful they should not create too many alarms in the following three months when nothing unusual was going on but they should certainly detect aero braking related activities.

The initial data pack consisted of one year of telemetry data of the VEX satellite (in the period 1st of March 2013 – 28th of February 2014), including over 15,000 parameters, called in the following nominal dataset. Moreover also log files and database files related to global mission data and orbit information data were provided, including parameters in the Spacecraft events set (SES), tele-commands (TC), parameters related to communication and data management system set (SCOS, Spacecraft Operating System) and parameters in the Auxiliary sets.

Moreover SATE also extended the analysis to the telemetry data of the XMM-Newton satellite (XMM) that were available from a preceding project.

IV. What is useful?

The first request SATE made to ESA was for a definition of “useful check” in the operational sense. This was by no means an easy question to answer. After several iterations the following formulation was made.

A useful check is automatically derived and applied, simple and can be quickly performed. It gives a good indication of significant change in behavior (statistically significant / different enough).

A not useful check is one that produces many false alarms.

Using this definition as a starting point SATE took a very pragmatic, engineering type approach to the problem. They defined algorithms to search for anything that could be classed as constant in the data. This could be simple features of the data such as the average or more exotic features such as the harmonic mean. The simple but powerful idea was that what was constant in the past should be constant in the future.

V. Features checkability

The purpose of the features checkability analysis is to identify the set of features and checks to be applied to those features that are potentially useful to detect novelties in the behavior of the parameters of a spacecraft, over the entire time period (i.e. always) or over multiple time periods (i.e. under certain conditions). A feature is a characteristic of the parameter time series that can be calculated over a given time window (e.g. mean, standard deviation, minimum, etc).

44 features for numerical parameters and 12 features for categorical parameters were defined by SATE, including simple features such as the mean and the standard deviation and more complex features such as FFT coefficients and features characterizing the sampling rate. The reason for checking features, instead of the original parameters, is that features characterise the parameter behaviour in a way that may allow identifying different novel behaviours in the parameter which may not be detected from the analysis of the original parameter time series (for example the coefficients of the FFT characterise the frequency content of the parameter time series).

The following types of checks were used, whereby it is highlighted that, in this context, the term “nominal” is not necessarily associated to the notion of “normal” or “fault-free”, rather it means “what observed so far”, which in most cases could be also “normal” but not as a general a priori rule.:

- 1) Domain check: this check includes two types of check, named Range check and Values check. These verify if the range or values of the feature change over time with respect to the reference ones, computed from a nominal dataset.
- 2) Variability check: this check verifies if the trend of the feature changes over time with respect to the reference one computed from a nominal dataset.
- 3) Distribution check: this check verifies if the distribution of the feature changes over time with respect to the reference one, computed from a nominal dataset.
- 4) Frequency check: this check verifies if the frequency content of the feature changes over time with respect to the reference one, computed from a nominal dataset.
- 5) Inter period check: this check verifies if the modes of a feature (among those with modal behaviour) change over time with respect to the reference ones, computed from a nominal dataset.
- 6) Few samples check: this check applies to the parameters that do not have enough samples to compute any feature. It consists in verifying if the parameter values change over time with respect to the reference ones, computed from a nominal dataset.

All checks, except the Few samples check, are applied to features time series, computed from the raw parameter time series. The Few samples check is applied to the original parameter time series, due to the fact that it is applied only to parameters that do not have enough samples to be analysed by the other checks. The usefulness of a feature is connected to its checkability which must be:

- 1) Significant: it provides valuable information on the behaviour of the feature (and related parameter) to which it is applied, allowing at least fault detection, i.e. the determination of the presence of a fault in a system.
- 2) Reliable: it does not generate too many false positives.

VI. Algorithms

SATE developed two algorithms: FETCH and Inter period analysis. They both are discussed in this section.

A. FETCH

FETCH stands for Feature Checkability. This algorithm aims at identifying the shortest time window (and consequent warning delay) beyond which Domain check, Trend check, Distribution check and Frequency check are useful (significant and reliable) for each feature, thereby identifying automatically the settings required to perform the check over the entire time period, e.g. Time Window Duration (TWD) for feature computation, reference values, thresholds, etc..

One of the main results found by this analysis is that features checkability over the entire time period is strongly related to the Time Window Duration (TWD) used for the computation of the feature or its characteristics (e.g. distribution, frequency).

Indeed the results showed that the same feature computed for the same parameter with different TWD may have different behaviors which lead to different checks to be applied.

For example, considering the parameter shown in Fig. 1 and the feature minimum, it is observed that using a TWD equal to 1 hour the feature shows a recognizable periodicity (Fig. 2) which allows performing a Frequency check, to verify that the frequency content remains the same; instead, using a TWD equal to 36 hours (Fig. 3), the feature has a constant behavior which allows applying a Domain check to verify that the values of the feature do not change over time. It is worth highlighting that these two checks are complementary and may be applied in parallel.

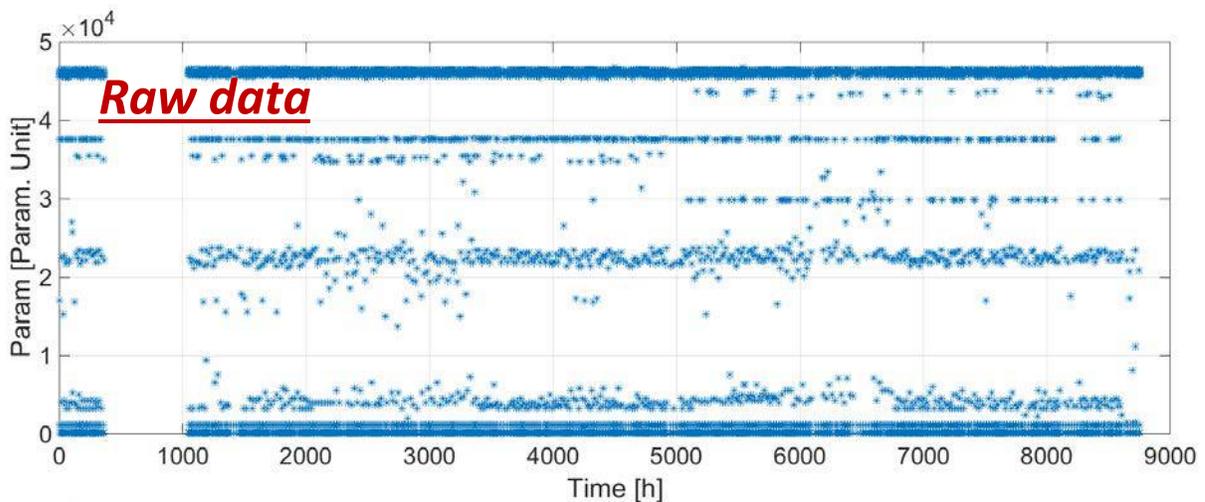


Figure 1. Example of one satellite telemetry parameter.

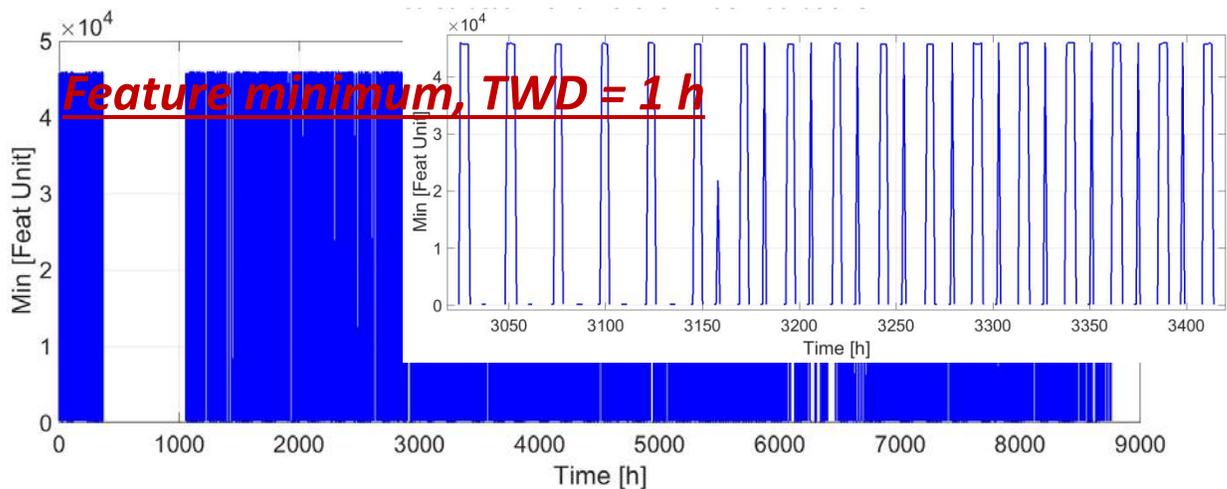


Figure 2. Feature minimum computed with TWD = 1h.

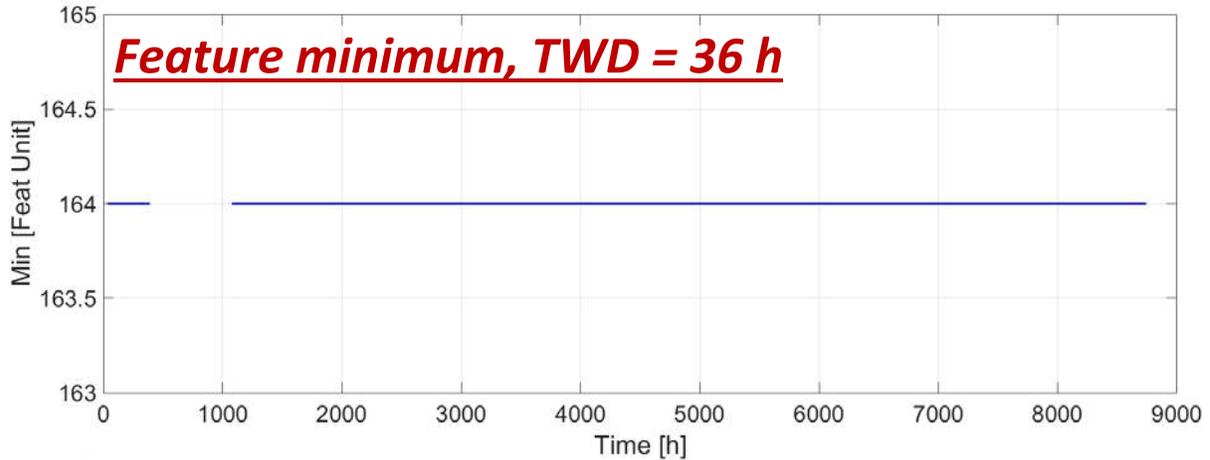


Figure 3. Feature minimum computed with TWD = 36h.

B. Inter period analysis

The Inter period analysis algorithm allows assessing the checkability of the features over multiple time periods by aiming at identifying the features with modal behavior and characterizing their modes which may be associated to (unknown) external factors that may influence the parameter behaviours. For these features the identified modes or patterns represent multiple “reference behaviours” that can be used to compare a new observation (i.e. new feature data) with suitable settings (e.g. reference values, thresholds, etc.). The methods identifies automatically features with modes and the settings required to perform the check. The time windows used for the evaluation of the feature behavior by the Inter period analysis is the same for all the features and may be determined on the basis of the output of the Frequency checkability analysis, as the most recurrent frequency.

Figure 4 shows an example of feature time series (feature Arithmetic mean of the parameter NTSA2002) showing different behaviours during the entire year period. In the figure, the colours are associated to as many patterns or modes that are exhibited by the feature and that are identified by the Inter period analysis.

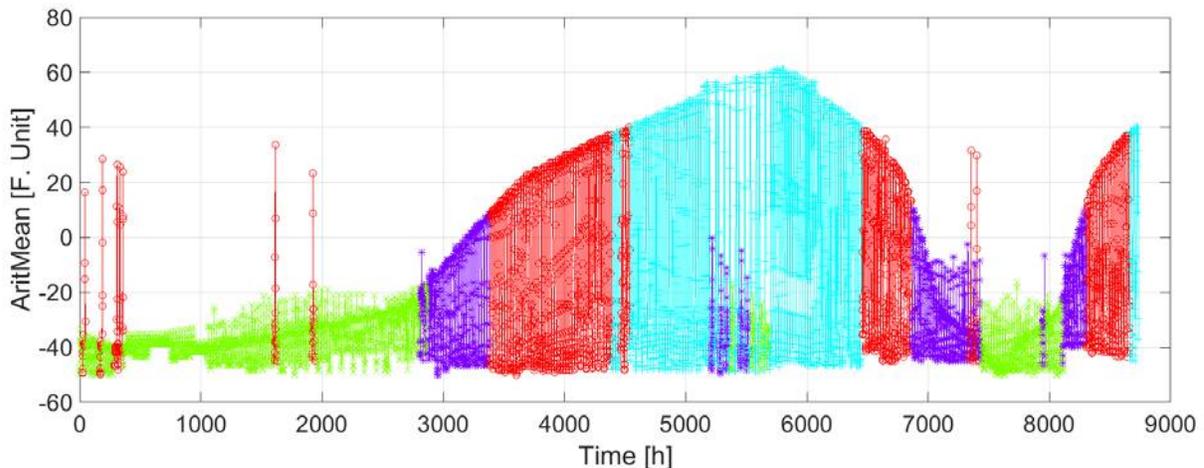


Figure 4. Example of a feature time series (Arithmetic mean of parameter NTSA2002) along the whole yearly history, with four different patterns (or modes) identified by the *Inter period* analysis.

FETCH and the Inter period analysis are consistent in selecting the appropriate checks to be applied to each feature of each parameter. It was found that all parameters in Venus Express and XMM are checkable by at least one of the six checks defined. The few samples check was applied to the parameters that were not checkable by the other five checks because they had a low number of samples. For both Venus Express and XMM, it is observed that

Domain check, Distribution check and Inter period check are the most recurrent ones, which can be applied to the majority of the checkable parameters.

VII. ESA Validation

For the validation exercise, historical data from the period 03/2013-03/2014 of the Venus Express mission was used (identical to the data provided to SATE for analysis). The algorithms proposed by SATE were implemented by ESOC in Java and then tested on telemetry data from the period 03/2014 -06/2014 (data that SATE did not have at that time). Java was chosen for two main reasons. It turned out that Java has a significant speed advantage compared to other languages usually used for statistical computing such as Python, Matlab or Stata which is of importance as runtime was a crucial aspect for feasibility. The second one was compatibility as most existing data-analysis tools in ESOC are written in Java. The telemetry data, which was stored as .csv-files on external drives, was read through a program and converted into java-arrays.

The choice of test-data was legitimate as the period 03/2014-06/2014 contained events of significance to spacecraft operators such as orbital control maneuvers, quadrature operations and the aero-breaking experiment, as well as periods that should be considered nominal. This allowed us to test whether the false alarm rate was reasonable and at the same time important operational events could be picked up automatically by the different approaches. The general idea was to count on a daily basis how many parameters failed the proposed checks i.e. the algorithm was declaring the behavior as novel. This was then cross checked with the Mission Report for that day to see if significant events had been reported by the Flight Control Team or not. This was also done in the opposite direction, i.e. when significant events were reported in the Mission Reports the algorithm results were checked to see if anything had been reported.

For the validation 10767 different parameters were identified, that were split into 5565 numerical parameters and 5202 categorical parameters. 35 numerical features and 12 categorical features were calculated over 10 different time window durations. This gave rise to more than half a million possible time series.

The first step of the validation was to generate all the necessary time-series from the raw data. Each parameter-feature pair was generated and the SATE algorithms used to generate the checks were validated. During the feature calculations it became clear that certain statistical features could not be used. Features such as skewness and coefficient of variation do not make sense on certain types of data for mathematical reasons, making it impossible to use the algorithms independently of the data. Other features were too easily impacted by sampling time and sampling irregularities resulting in unreliable numbers and generating obvious false alarms. Removing these feature-parameter pairs allowed us to significantly reduce the number of false alarms later on.

VIII. Only Strict Checks Worked in the Validation

The SATE results had shown that the Domain check and the Distribution check were the most applicable for the test data hence we decided to validate these two types of check.

However, this does not imply that the other types of checks are not useful. Indeed, results obtained in the work indicate that all checks, including Trend and Frequency checks, are capable of detecting different types of novelties. The reason for the low number of novelties detected by Trend and Frequency checks is that a lower number of parameters and features are checkable by those checks, i.e. have a trend or frequency content that can be checked in a reliable way. However, when these characteristics are present in a feature time series, it is deemed useful to check them by Trend or Frequency check.

Several check criteria for feature time-series for the Domain check were investigated. The results showed that although the check criteria could be varied for each parameter to make it less or more strict we found that any other criteria besides the strictest available produced far too many false alarms. For example, a “sufficiently” small coefficient of variation that would make sense for all parameters could not be chosen. The value had little to do with the data itself and was much more an engineering decision – something that would break one of the golden rules.

Although disappointed, the idea of only applying very strict checks did bear fruit. We decided to take the idea of looking for only constant parameters or features to the limit. Our checks were only performed on features consisting of less than three different values, or raw time series containing less than ten different values. Recalling that the SATE approach was to define algorithms to search for anything that could be classed as constant in the data, the most striking observation during the validation was that the vast majority of telemetry data could indeed be described in some constant way using this strict approach. It is worth noticing that with these checks roughly 95% of telemetry data from Venus Express could be checked, out of which 81% parameters were constant and 7% of parameters had less than 10 different values (Fig. 5). This raised the question on whether sophisticated checking algorithms were needed at all to cover the vast majority of the data.

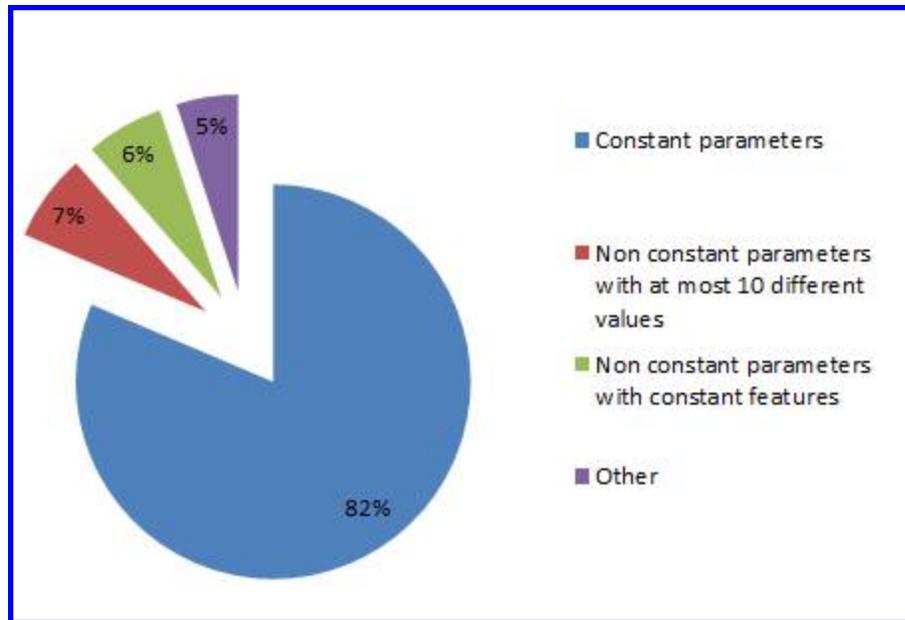


Figure 5. Telemetry parameters characterization of the Venus Express satellite.

Using these strict criteria the tests were then run on the future time period. The results are shown in Fig. 6 and are quite impressive. The aero-breaking experiment was clearly picked up on Day 80 and also orbital control manoeuvres on Day 16 and 32. Furthermore the feature time-series were able to pick up a loss of telemetry on Day 66 as this strongly impacted the feature calculation and the features characterising the sampling rate (e.g. average and standard deviation of the sampling rate). We realized that this could be used to our advantage as the algorithms were naturally sensitive to any change in the sampling rate providing a latent check on this.

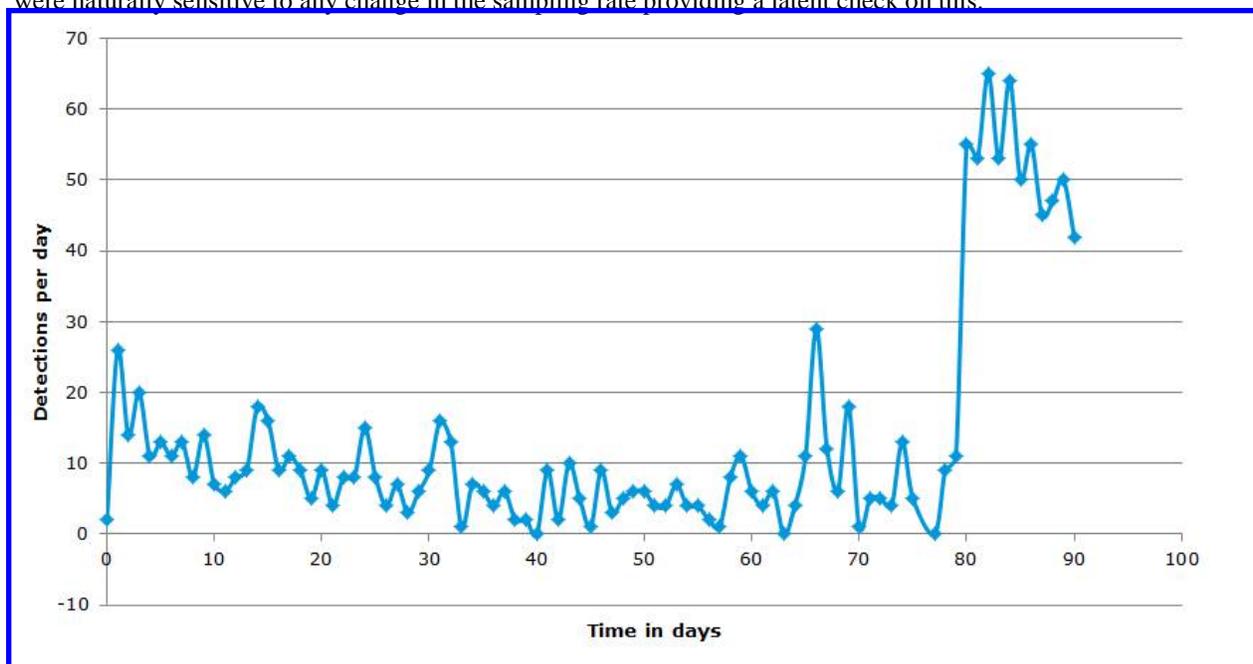


Figure 6. Number of novelties per day in the three-months validation period (90 days), considering only features consisting of less than three different values, or raw time series containing less than ten different values, accounting for 95% of telemetry data from Venus Express.

The results of the Distribution check did not give good results during the validation checks. Again the criteria seemed not meaningful enough to give information whether a time-series was checkable. Even with varying

boundaries this check produced too many false alarms. More work is needed in this area to make it meaningful. However, in some examples the Distribution check detects some novelties earlier than Domain and Inter period check.

We concluded during the validation phase that searching for constant parameters, or parameters or features that only took a limited number of values then checking for new values in the future gave real results for the eligible parameters. These sorts of checks cover a surprising amount of the telemetry dataset. Moreover it turned out that this is not a phenomenon of the Venus Express mission, it is also true for every other mission for which telemetry data was available. An advantage was that the computational effort involved in this sort of checking is extremely small and therefore fast. The disadvantage was that not all parameters can be checked like this and that checks in real time is not always possible – as some of the features have to be calculated over certain time windows which last up to a week, depending on the type of check and feature according to the automatic analysis of the parameters by FETCH. This aspect, however, does not prevent the application of the methods as “online” algorithms as long as the results delay can be accepted.

IX. SATE react to the validation results

The validation exercise highlighted a particular problem of the checking process i.e. it was essentially threshold-based and therefore binary i.e. either the check passed or failed. We concluded that to avoid many false alarms we needed to set strict limits with the consequence that only sudden major changes could be reliably detected. SATE approached the problem from a different angle, saying that what was needed was something else to measure how novel the new behavior was compared to the past history. They called this a Symptomatic Variable (SV), referred to later also as priority score. This variable essentially removed the binary nature of the check allowing degrees of novelty to be displayed.

Priority scores are synthetic measures to quantify how different the behaviors of the parameters are in two different time periods. The greater the difference, the higher the priority score. This allows the generation of priority lists, i.e. lists of parameters ordered according to the priority scores, indicating which parameters should be further analyzed first.

Figure 7 shows an example of priority score computation for a predefined check. When the new observation falls within the check range a negative exceedance is computed which is associated to a normal behavior of the system (null priority score). When the new observation falls out of the check range then a positive exceedance is computed which is associated to a novel behavior and its degree of novelty is quantified as the ratio between this exceedance and the check range.

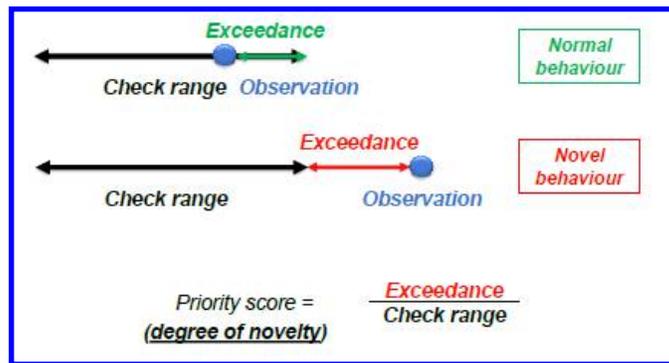


Figure 7. Example of priority score.

When the extra dimension provided by the priority scores were calculated for the checks defined in the previous sections then a much richer picture emerged. This created a synthesis of the overall status of the spacecraft over time and easily allowed identifying days where important operational events took place and even in which subsystem. Cross correlations were also much easier to identify in this view.

SATE applied all the checks defined using the nominal dataset to a new dataset of the VEX satellite used for validation. This dataset, called in the following validation dataset, started with the three months used by ESA during its validation activities (03/2014-06/2014) and extended up to the up to the 31st January 2015, including a total of 11 months. However, stable telemetry data are present only up to the first days of December 2014, thus including, about 9 months data, starting from the 1st of March 2014.

As already said, since Venus Express far exceeded its planned life, the spacecraft was tasked with conducting a series of aero-braking manoeuvres. The major aero-braking campaign was conducted from the 18th of June to the

11th of July. Since these manoeuvres were not performed during the period covered by the nominal dataset initially provided by ESA, a large number of novelties is expected to be detected by checks during the comparison period (after the first 60-80 days). After the aero-braking campaign, the orbit was raised to try and return back to the mission orbit phase. Since then, the mission continued in reduced scientific phase. This implies that the number of samples (telemetry) for the parameters associated to the scientific instruments are reduced.

The change to the reduced scientific phase is expected to result in a novel behaviour of the features associated to the sample rate of the parameter time series.

Figure 8 synthesises the main events occurred in the new time period under analysis, used for validation.

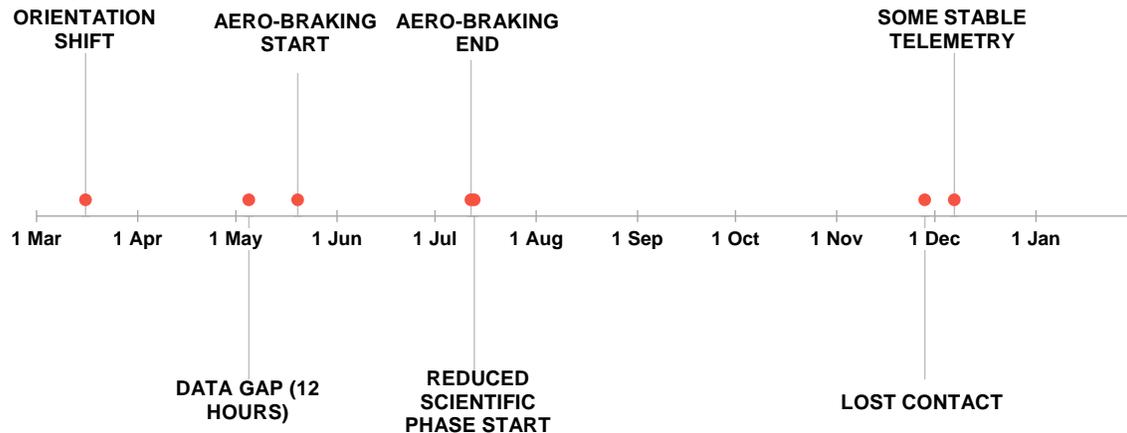


Figure 8. Time line of some major events regarding the VEX satellite, occurred between the 1st of March 2014 and the 31st of January 2015, used for validation.

Figure 9 shows the three dimensional distribution of the Symptomatic Variable (or priority score) computed by the Domain check for all numerical features of the Venus Express satellite. It is possible to see that in the first quarter the number of novelties are less and characterized by a lower Symptomatic Variable with respect to the next quarters in which the number of novelties increases a lot and are also characterized by higher Symptomatic Variable. As already anticipated during these quarters a series of aero-braking maneuvers were performed and a reduced scientific phase was started, lasting until the end of the validation dataset.

This was considered to be a significant result of the study allowing an ESOC flight control team a completely different view on the spacecraft status in an instance.

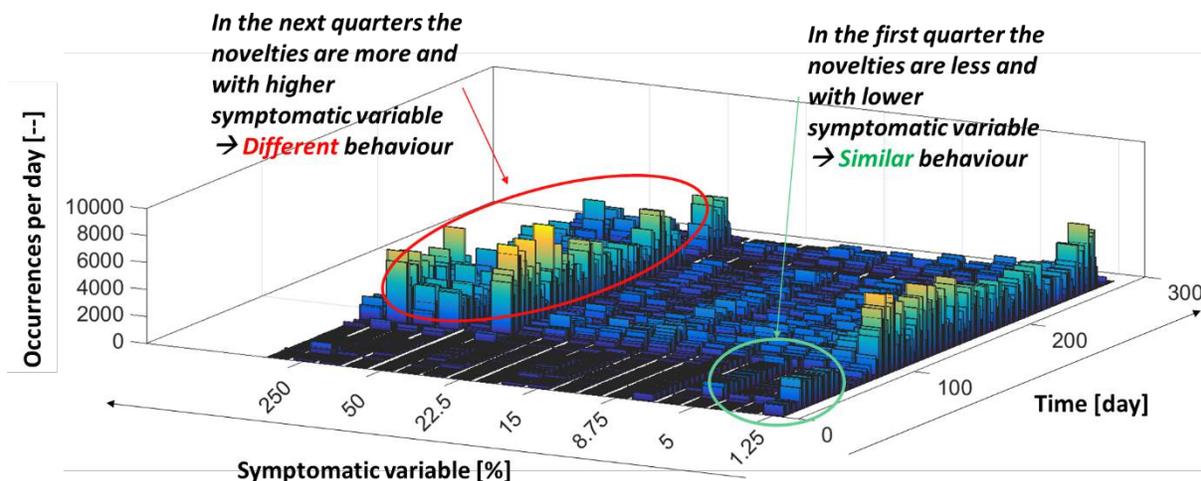


Figure 9. Number of features with novelties, detected by *Domain* check for numerical parameters in the VEX validation dataset, distributed according to the symptomatic variable (priority score) and to the time period analyzed.

Analysing in greater detail the number of parameters with novelties and the value of the Symptomatic Variable allows a better understanding of the results and of their possible impact in the improvement of the efficiency of ground operations. Figure 10 shows the number of numerical parameters with novelties in the validation dataset detected by the Domain check selecting only the parameters characterized by novelties with a SV greater than a given threshold. It is possible to see that:

- During the initial part (day 1-80), when the satellite is expected to behave normally, the number of parameters with novelties is:
 - i. 100, using a threshold for the SV of 10% (green area of the plot)
 - ii. 25 using a threshold for the SV of 50% or 100% (blue and red areas of the plot)

Therefore in this example the Flight Control Engineers (FCEs) should check only about 100 parameters with a SV threshold of 10% or even just 25 with a threshold of 100%, per day.

- During the second part (day 80-290) the number of parameters showing novelties remains high even if only the parameters characterized by novelties with high SV are considered. Therefore, this confirms the results commented above that the parameters showing novelties in the second part are also characterized by high SV.

This is in line with the expectations since it is known that the VEX satellite in this period had a different behaviour with respect to the reference period.

In any case it is highlighted again that the SV allows generating a priority list indicating which parameters should be verified first by the Flight Control Engineer (FCE), considering also his/her time availability. Therefore the FCE is not expected to check all the newly listed parameters.

Indeed a short list with the first N parameters (where N is a user-defined parameter) could be provided, as the ESA's Novelty Detection already does.

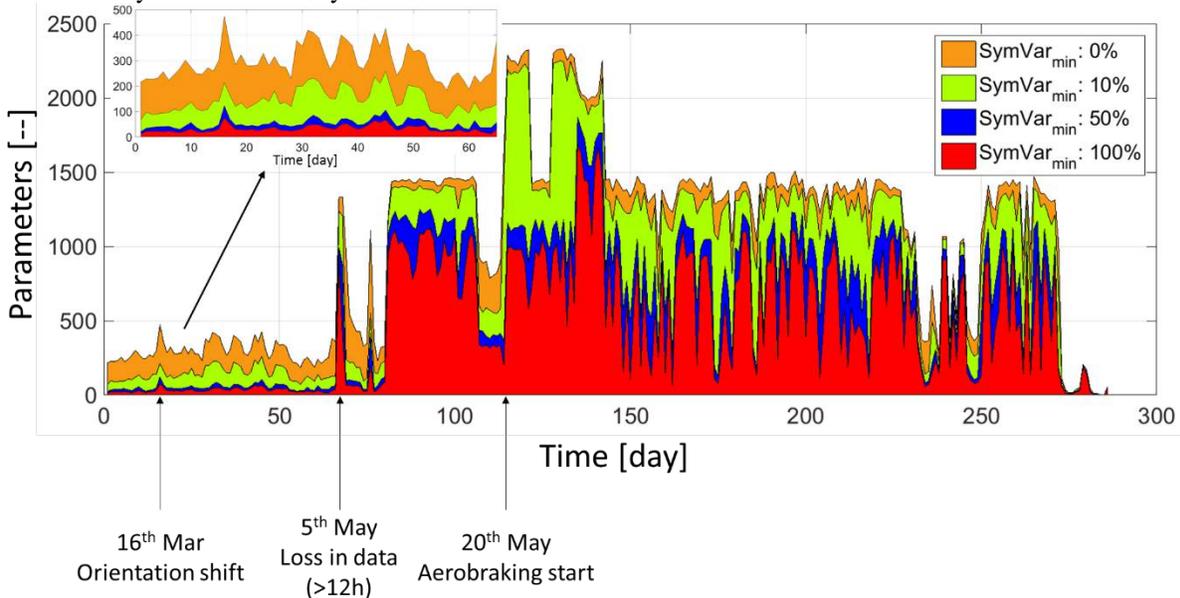


Figure 10. Number of numerical parameters with novelties detected by *Domain* check in the VEX validation dataset, in each day, grouped for different values of symptomatic variables.

In order to further focus the FCE attention to significant novelties, the number of parameters showing novelties on a K days-basis was analysed, to remove from the parameter counting those that show novelties for several consecutive days. This analysis is done by counting only one time a parameter showing a novelty at the jth day and by not considering the novelties occurring in the same parameter in the following K days.

For example, if a week-basis is adopted, K is equal to 7. Then if a parameter showed novelties:

- From day 1 to day 9 and
- In Days 18 and 19

Only days 1, 8 and 18 account for 1 novelty.

Figure 11 shows the example described above comparing the normal counting in which a parameter is counted for each day in which a novelty is detected (upper figure) and the week-basis counting ($K=7$) in which the novelties occurring in the following 6 days ($K-1$) after that a novelty occurred are not counted (lower figure).

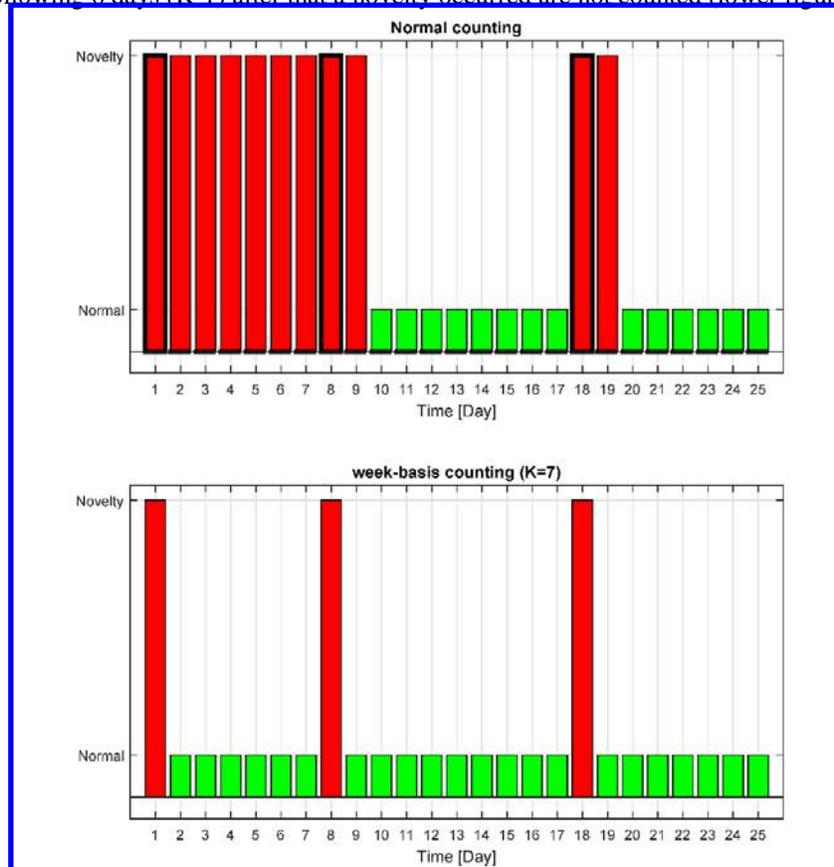


Figure 11. K-days analysis - Comparison between normal (upper) and week-based (lower) counting of novelties.

The K days analysis was performed on the validation dataset using the following different time intervals (i.e. different K):

- $K = 7$ (1 week, see Figure 12)
- $K = 31$ (1 month, see Figure 13)
- $K = 365$ (1 year, see Figure 14)

The results showed that:

- Clear recurrent peaks are visible, almost every K days (in the week and month counting, Figure 12 and Fig. 13 respectively).

This means that the parameters showing novelties are more or less always the same ones.

- This is further confirmed by the year-basis analysis (Fig. 14) in which only three peaks are detected on the 16th day (16th of March 2014 corresponding to the orientation shift) the 67th day (5th of May 2014 corresponding to loss of data) and the 114th day (22nd of June 2014 during the aero-braking manoeuvre) whereas *on the days after the peaks there are almost zero novelties*.

Therefore these results proved that, if the FCE focuses only on the “first novelty” occurring on a parameter, the number of parameters to be checked significantly decreases. This would avoid checking parameters that have been already checked in the recent past by the FCE and that can be supposed to having been managed.

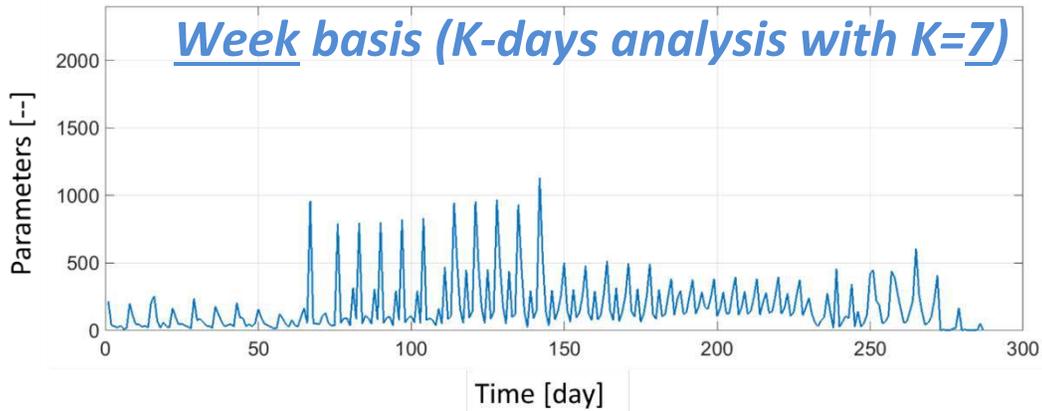


Figure 12. Number of numerical parameters with novelties detected by *Domain* check in the VEX validation dataset on a week basis (K-days analysis with K=7).

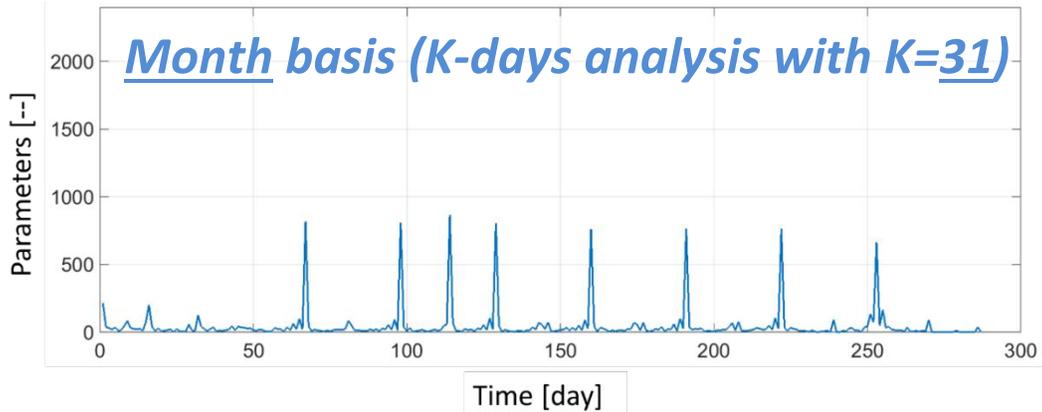


Figure 13. Number of numerical parameters with novelties detected by *Domain* check in the VEX validation dataset on a month basis (K-days analysis with K=31).

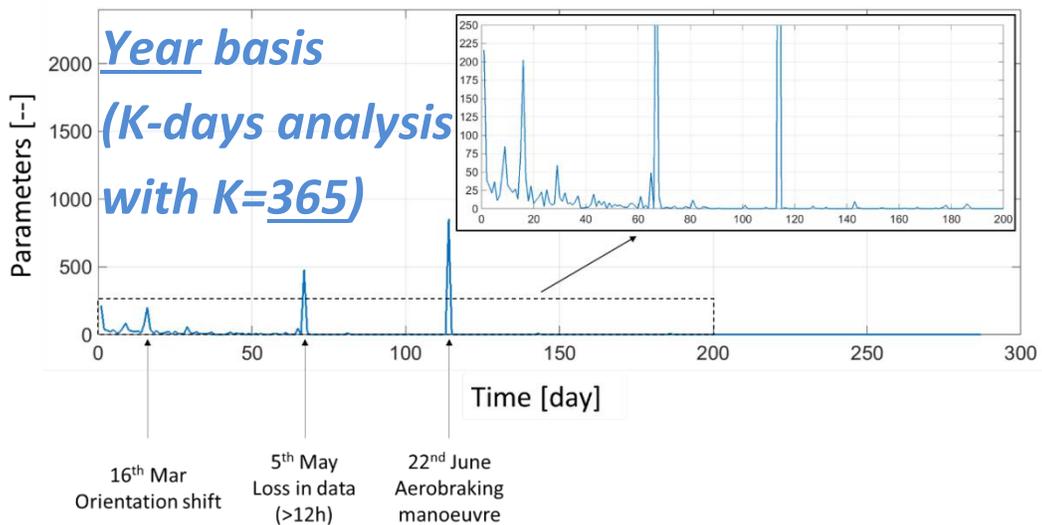


Figure 14. Number of numerical parameters with novelties detected by *Domain* check in the VEX validation dataset on a year basis (K-days analysis with K=365).

X. Conclusions

To conclude we have shown that it is possible to dramatically improve present day checking of spacecraft telemetry. This is because the majority of the data can be described as quasi-static i.e. there is some way to describe it's properties as a limited set of values. SATE produced algorithms to try to find these quasi static properties in each telemetry parameter in order to make it checkable. If the parameter was not checkable in its raw state then they calculated features and adjusted the time window over which the feature was calculated until they did find something that was quasi static and therefore checkable. Using this strategy they successfully managed to produce checks for all of the parameters in the VEX data (except one numerical parameter) and all of the XMM satellite. The advantages of this approach are its ease of implementation and low processing power cost.

The ESA validation tests did show a disadvantage, namely the number of parameters failing the checks in the future time period. In response SATE showed how it was possible to provide useful indications to Flight Control Engineers about which parameters should be analysed first from a large list of failed checks. Indeed we used a rule that no more than 10 new novelties per day should be produced unless there was a major mission change. The symptomatic variable, which is a quantity that measures the degree of novelty in the time series analysed, was able to do this and moreover group views of different telemetry parameter's symptomatic variables provided a synthesis of the overall status of the spacecraft over time that is completely new for us. Initial proof of concept suggests that it allowing quick identification of events that have impacted multiple telemetry parameters simultaneously for instance.

XI. References

¹Martinez-Heras, J.A., Donati, A., Kirsch, M. G. F., Schmidt, F.; "New Telemetry Monitoring Paradigm with Novelty Detection"; In SpaceOps 2012 Conference, Stockholm, Sweden, June 11 – 15, 2012.

²Martinez-Heras, J., Donati, A., "Enhanced Telemetry Monitoring with Novelty Detection" AI Magazine, Vol. 35, No. 4, 2014, pp. 37, 46.

³Jacopo Biancat, Chiara Brighenti, Attilio Brighenti, José-Antonio Martínez-Heras, Alessandro Donati, David Evans, "Priority Scores based on novelty detection to improve the efficiency of ground operations", Big Data from Space, Tenerife, March 2016