ESA STUDY CONTRACT REPORT

No ESA Study Contract Report will be accepted unless this sheet is inserted at the beginning of each volume of the Report.

ESA Contract No: 4000112223/14/F/MOS	SUBJECT: Automatic Spacecraft Status Characterisation by data mining mission history - Executive Summary	CONTRACTOR: KU Leuven
* ESA CR()No:	No. of Volumes: 1 This is Volume No:1	CONTRACTOR'S REFERENCE: KUL-IVS-ASSC-RP-006

ABSTRACT:

This study investigates data mining techniques useful for the health monitoring of spacecraft, focusing on (1) features derived from telemetry parameter time series; (2) derived parameters from auxiliary data sources; (3) Kernel-density estimate (KDE) based outlier detection; (4) KDE based data density change detection; (5) Poincaré plot analysis; (6) command effect analysis

Names of authors:
Wim De Meester, Joris De Ridder, Rik Huygen, Sara Regibo, Pierre Royer, Bart Vandenbussche

** NAME OF ESA STUDY MANAGER: David Evans DIV: Advanced Operations Concepts Office DIRECTORATE: Human Spaceflight and Operations	** ESA BUDGET HEADING:

* Sections to be completed by ESA
** Information to be provided by ESA Study Manager

Automatic Spacecraft Status Characterisation by data mining mission history Executive Summary

Issue 1.0

25-Nov-15

Wim De Meester, Joris De Ridder, Rik Huygen, Sara Regibo, Pierre Royer, Bart Vandenbussche

EUROPEAN SPACE AGENCY
CONTRACT REPORT

The work described in this report was done under ESA contract.

Responsibility for the contents resides in the author or organisation that prepared it.

This document contains no Proprietary Information

Institute of Astronomy KU Leuven
Celestijnenlaan 200D – bus 2401
3001 Leuven
Belgium

Table of contents

1 Introduction

The present document provides the Executive Summary of the ESA GSP study Contract No. 4000112223/14/F/MOS "Automatic spacecraft status characterisation by data mining mission history"

Over the last decades, the increasing downlink bandwidth and the processing power available on Space Platforms has lead to an escalation of the number of telemetry measures available to spacecraft controllers to monitor the health and safety of spacecraft. The number of housekeeping parameters on large platforms is typically in the order of tens of thousands. While providing a wealth of diagnostic information, these vast numbers exceed the possibilities of the human brain to oversee all telemetry measures. Automatic checks and computer-aided data analysis are needed to help spacecraft operators and operation engineers keep the overview.

Traditionally, mission control systems monitor incoming telemetry measures against pre-defined soft and hard limits. Spacecraft operators are warned in real-time and in out-of-limit logs when telemetry parameters go out of limit. The trend towards more autonomous spacecraft also has lead to the use of event packets on board. Besides triggering autonomous recovery actions on board, these event packets are also the first warning flags for spacecraft operators to assess the health and status of the spacecraft when acquiring ground contact. However, some spacecraft produce a vast number of event packets, which also make it difficult to oversee them.

More advanced data mining techniques can assist spacecraft controllers in the three core aspects of telemetry monitoring. The first efforts to provide more powerful statistical tools for novelty detection in spacecraft telemetry have resulted in the first data mining systems available at ESOC integrated with the MUST telemetry archive. This study further elaborates on this activity.

The study has been organised in different activities.

The ESA provided reference dataset for validation of the developed methods was extracted into a convenient **reference data archive** with access facilities to raw data, and catalogues of parameter properties.

The next activity encompassed the identification and calculation of **generic telemetry parameter features.** Not only spacecraft telemetry parameter time series are available to describe the behaviour of the space segment. Auxiliary data sources are available, for example command history files, orbit, attitude history, event logs, etc. In Section 3 we report on the study activity on **features derived from auxiliary data.**

Different study activities looked at complementary techniques to **discover novelty** in time series of raw telemetry parameter measures and features. A first novelty discovery study activity provides a robust **density based single outlier detection method**, in which we look for single parameter readings that are significantly out of the nominal distribution of the values. The method is based on kernel density estimates, and is learning the nominal data distributions in the course of the mission, eventually learning that single outliers that happened before are no novelty. The second novelty discovery study activity is also based on the kernel density estimate. The novelty looked for now are **density changes**. The third novelty discovery study focussed on using **Poincaré plots** to discover novelty. Poincaré plots can reduce the apparent complexity of a signal to just a few data points. Especially for enumerated data series this provides a powerful detection of e.g. state transitions that have not been seen before. Also on continuous signals Poincaré plots provide information on the way the data changes.

The last study activity aimed at learning correlations between commands and changes in telemetry values or their features. In the **Command Effect Analysis** the typical footprint of all command executions is determined by finding the parameters and parameter features that consistently show a similar before/after-command-change.

2 Telemetry features

The goal of this activity was to make an inventory of features that hold potential for increased sensitivity of knowledge discovery in time series. Features should typically provide complementary sensitivity to the raw time series for novelty detection algorithms. Different features were identified for continuous measures (e.g. Temperatures) and enumerated features (e.g. status words).

The feature time series are typically calculated from data points surrounding the original time series data points, e.g. a detrended feature by subtracting the mean of the time series over a time window around the data point. Both moving time windows of fixed duration and incremental windows were considered. Feature calculation algorithms were implemented and applied on the Venus Express reference data set. The following five length of windows were chosen: 3 minutes, 1 hour, 1 day, 10 days and “growing”. By growing window, we mean a window that is not fixed in length and sliding over the time series, but always starting at the same time as the time series, and growing with time at its other end. So, at every moment in time, the growing window encompasses the entire time series. This is a practical feature window concept for novelty detection in an operational mission control system, where data beyond the last sample is obviously not available.

We identified a list of features useful for continuous time series, as well as a list of features useful for enumerated or status time series.

Based on the features calculated from the Venus Express telemetry parameter time series, a feature sensitivity matrix was constructed. As an indication of how useful a feature is, the density based outlier detection algorithm (Section 4) was applied on all raw time series and calculated features over different time windows. The feature sensitivity matrix lists the number of operational days with at least one outlier detected per parameter/feature combination.

3 Derived parameters from auxiliary data sources

Auxiliary data sources (spacecraft event files, telecommand history files, orbit and attitude data, ...) were analysed to determine meaningful derived parameters, useful for novelty detection and command-parameter correlation analysis. The derived parameters were extracted from the auxiliary data corresponding to the Venus Express telemetry parameter dataset.

4 Density-based novelty detection: single outliers

Our goal is to identify anomalies that we define as telemetry behaviour that is so different from its historical behaviour that it raises suspicion that it is caused by a different mechanism. In this Section we focus on *single outlier detection*. One of the most simple outlier detection methods is σ-clipping. σ-clipping works well when the overall distribution of the telemetry data is not too different from Gaussian. For very skewed or multimodal distributions (which is typical for spacecraft telemetry) σ-clipping often cannot give the desired results.

We propose a non-parametric data-centric approach to estimate the probability density function (PDF) of the time series without making any assumption about it. One of the easiest ways to estimate the PDF is by making a histogram of the measured signal. However, alternatives exist that have better asymptotic properties, in the sense that they exhibit a lower bias and a lower integrated squared deviation from the true underlying distribution compared to histograms. The best-known non-parametric method is the Kernel Density Estimate (KDE). The KDE is calculated by summing the convolutions of every data point in the time series with a kernel (e.g. Gaussian) in the time domain. The KDE can then be used to calculate data intervals that are outside 'nominal' intervals that correspond to the larger fraction of the density estimate surface. This process is illustrated in Figure 1.

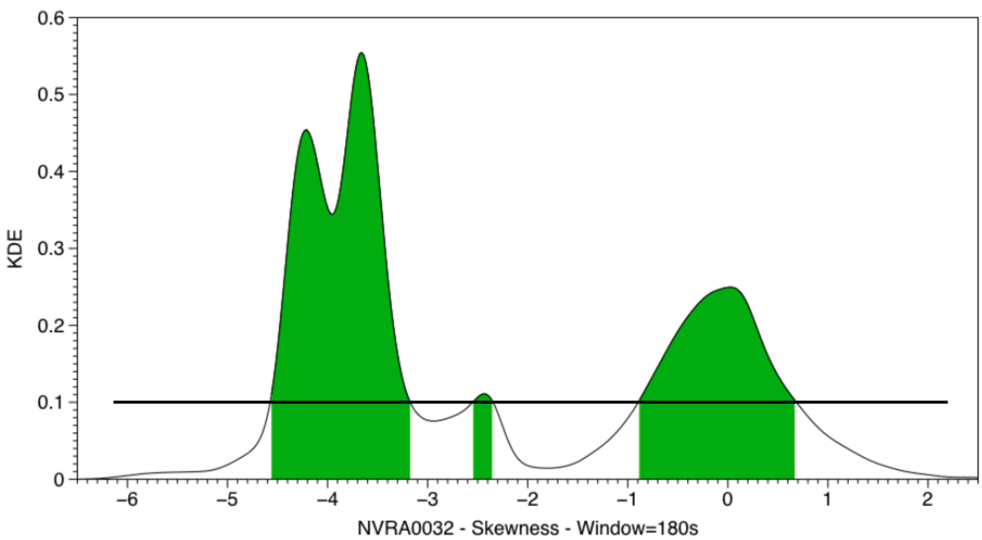

Figure 1: Illustration of how the nominal intervals are computed in practice. A horizontal line is gradually lowered until the sum of the green areas equals a predefined value 1-α. This defines a set of intervals (green colour) in which the values are considered nominal. All values outside these intervals are considered anomalous, because they occur extremely rarely (a fraction α of the time).

The study also investigated whether the KDE could be applied to the case of multiple telemetry parameters. The rationale behind this question is that two highly correlated parameters may show broad 1D KDEs for each of the parameters individually, while the 2D joint-KDE may actually cover a relative small volume, which offers a more stringent way to identify outliers. It turns out that this is indeed possible and multivariate kernel density estimates have a potential for more detailed anomaly detection.

The KDE outlier detection was validated against the Venus express mission history telemetry. An example is shown in Figure 2.

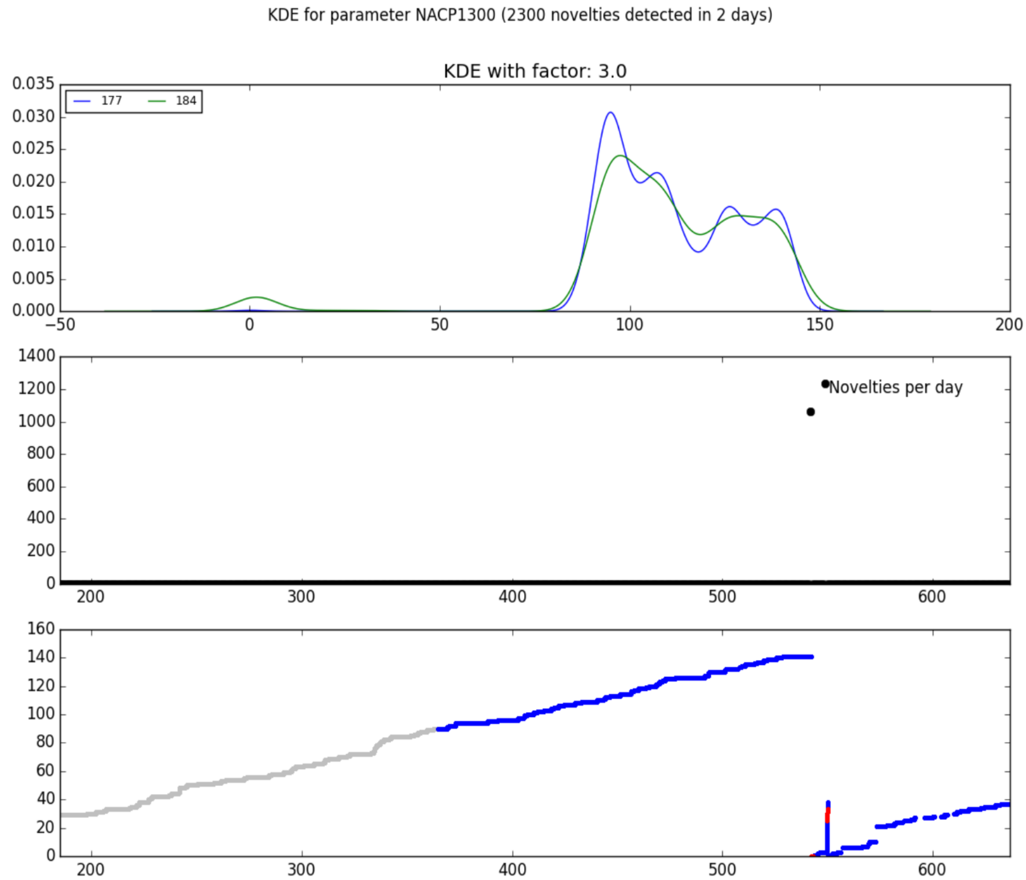

Figure 2: Illustration of the KDE outlier detection applied on a Venus Express telemetry parameter time series

5 Density-based novelty detection: density changes

5.1 KDE change detection

In the previous Section we assumed a fixed KDE, and use this distribution to identify outliers. In practice the KDE of a telemetry series is often not constant at all, and varies during the mission lifetime. In what follows we develop methods to detect when a KDE is changing ”abnormally” much. The idea is to split up the telemetry series in (possibly overlapping) windows, and provide a test to detect if the KDE in the current window deviates significantly from the KDEs observed in the previous window(s).

For each of the previous windows the KDE is computed and stored. The KDE of the current window is then compared with these previous KDEs: each time an anomaly measure is computed, and if this measure exceeds a pre-set threshold, the current window is flagged as anomalous. A last step involves the filtering of telemetry parameters. Some parameters may flag too often to be practical for an operator, and these are consequently not shown in order not to clutter the anomaly detection system.

In the following sections we assess three different methods to determine whether the KDE of the current window is anomalous w.r.t. the KDEs of the previous windows.

5.2 Detecting KDE changes using confidence bands

The idea behind this approach is to define an uncertainty ”band” around the KDE of the previous window, and check whether the KDE of this window falls within this band. If not, an anomaly is flagged. To illustrate this concept, we show in Figure 3 a KDE together with its 95% confidence band.

An example is shown in Figure 4. We applied the confidence band KDE variation detection method on the Venus Express telemetry parameter NAAD0336. We assessed whether the KDE of this week changed significantly with the KDE of last week, but checking whether the former is completely within the $\alpha = 10^{-5}$ confidence band of the latter. Windows of telemetry values plotted in blue are nominal, while those plotted in red are considered anomalous because the KDE changed too much. Tested on several types of parameters, the method turns out to be very sensitive, in the sense that even KDE variations that would be considered relatively mild when checking them by eye, are detected as anomalous. The great sensitivity is at the same time a strength as well as a weakness. For telemetry parameters for which the KDE ought to be very stable, we recommend exactly this method to monitor the parameter. For those parameters for which the KDE tends to vary quite a bit, we prefer other methods.

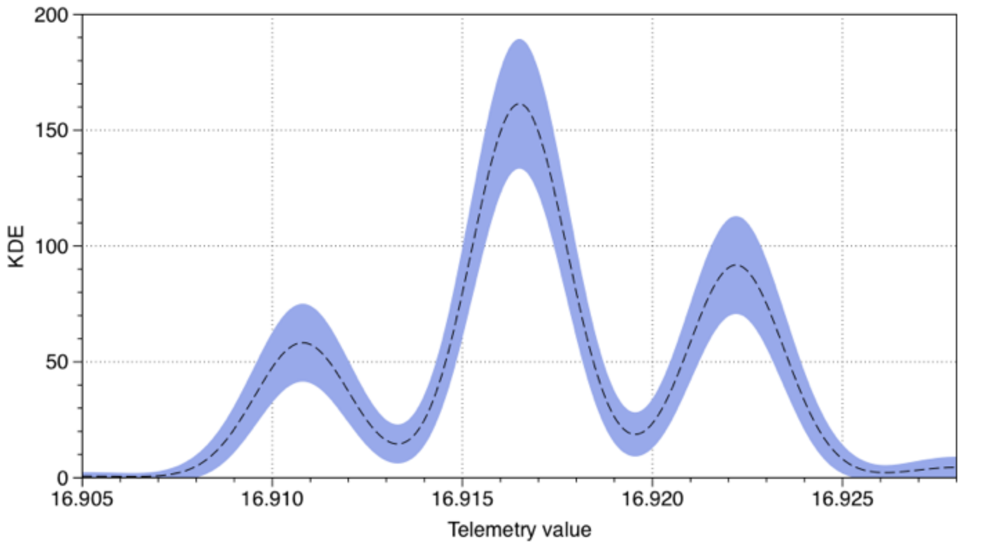

Figure 3: The dashed line is a KDE estimate of a telemetry parameter, while the blue band denotes the estimated 95% confidence band.

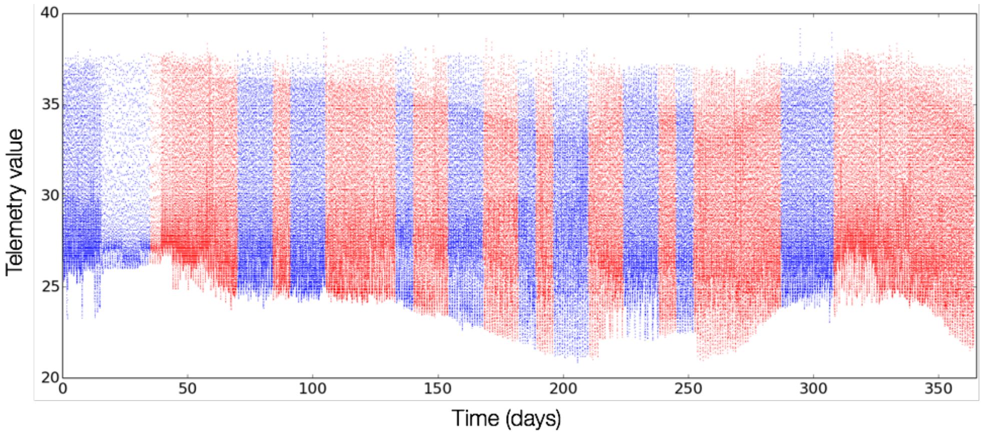

Figure 4: Time series of Venus Express parameter NAAD0336. The time series was partitioned in blocks of 1 week, and the KDE of each week was compared with the KDE of the previous week using the $\alpha = 10^{-5}$ confidence bands as a significance indicator. Blue weeks are considered nominal, red weeks anomalous. Clearly, the method is extremely sensitive to KDE changes, and can thus only be used for those parameters that ought to be very stable in time.

5.3 Detecting KDE changes using RMS bands

The second method we assessed to test for KDE variations, is based on establishing an "uncertainty band" using the observed KDE variations during the windows before the current one. First, we split up the time series in time windows. Given the first N windows, we compute the average KDE and the corresponding variance. In Figure 5 we show for a Venus Express telemetry parameter the KDEs of the different 7 day windows, and in red the (symmetric) RMS variation around the mean KDE. It is this RMS band that will be used to assess the KDE of a particular window. First, the mean and the RMS band are computed using the KDEs of all earlier windows. Then, it is checked whether the KDE of the current window falls completely within (say) 5 times the RMS band. If so, the window is considered nominal, otherwise anomalous.

We applied this method (using 5σ) on Venus Express telemetry parameter NTTG3001, for which we show the results in Figure 6. Apart from the obvious outlying intervals (coloured red in the top panel) it also detected the less obvious change around *t=20* d, where the range suddenly changed as plotted in the bottom panel. After a while this new range becomes the new "nominal" and the method no longer flags it.

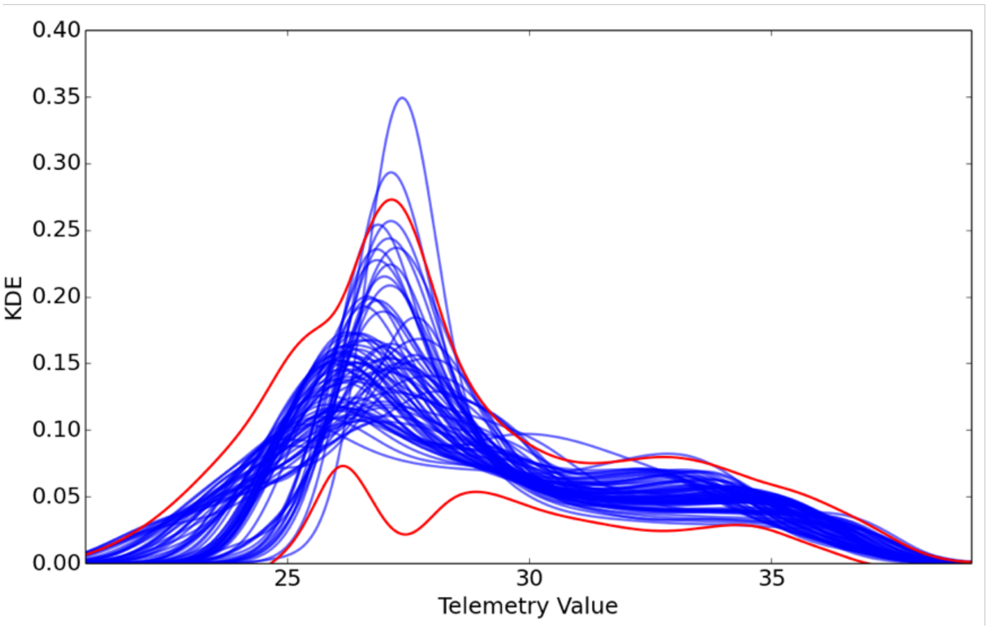

Figure 5: The 2.5σ RMS was computed for each telemetry value which results in an "uncertainty" band around the mean, delimited by the red curves, which can then be used to assess whether the KDE of a subsequent week deviates abnormally from what is previously observed.

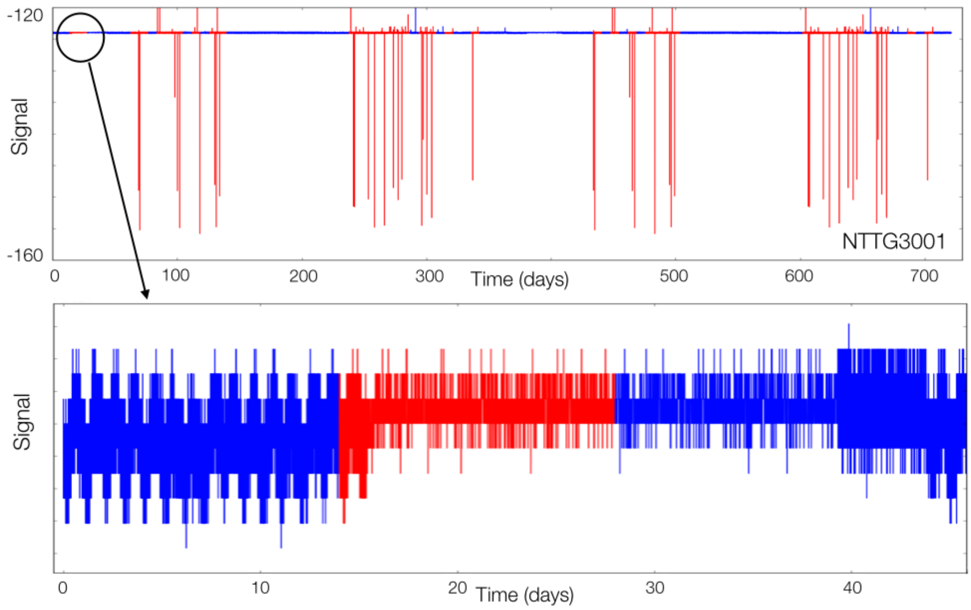

Figure 6: RMS band anomaly detection for telemetry parameter NTTG3001. Apart from detecting the more obvious outlying intervals (in red in the top panel), it also detects changes that the eye may not have immediately caught. This is illustrated in the bottom panel where a change in range is detected. After a while the smaller range becomes the new "nominal" and is no longer flagged.

5.4 Detecting KDE changes using area changes

The third method we assessed to test for KDE variations, is using the L_1 distance between two KDEs. First, we split up again the time series into time windows. We then calculate the L_1-norm (i.e. the integrated area difference) between the different KDEs, as illustrated in Figure 7. We compare this norm against a pre-set threshold, to determine if the current window of telemetry points is too different from what was observed in the past. If so, this window is flagged as anomalous.

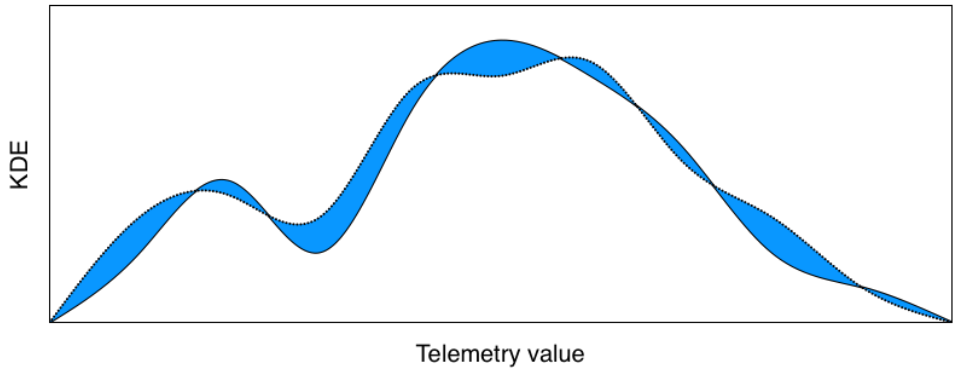

Figure 7: A schematic illustration of area difference $||f_n-f_k||_1$ between two KDEs $f_n(x)$ (solid line) and $f_k(x)$ (dotted line). Because the KDEs are normalised in area, also the area difference is bounded between 0 and 1.

To illustrate anomaly detection using the L_1-norm, we applied the algorithm to Venus Express telemetry parameter NAAD0431. The result can be seen in Figure 8, where we highlighted those days that showed an anomaly in red. The L_1-norm seems to be less sensitive to KDE changes than the RMS bands method, leading to less anomaly flags.

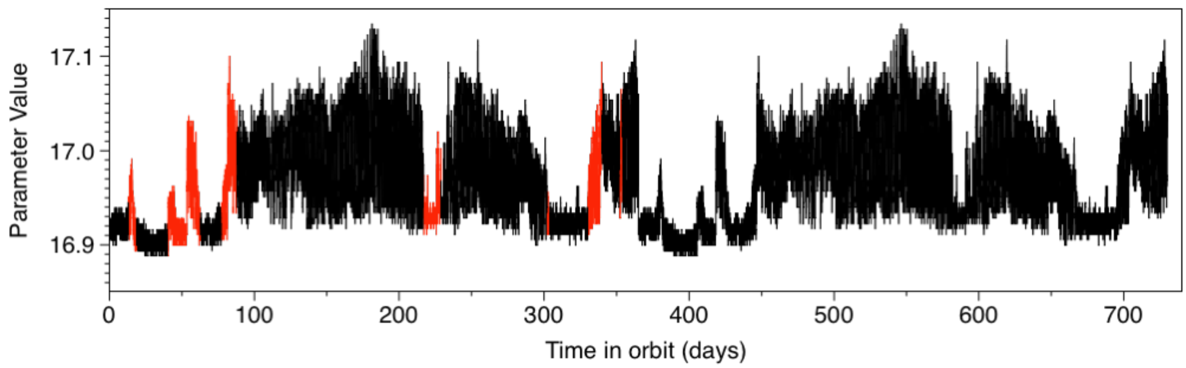

Figure 8: Anomaly detections for telemetry parameter NAAD0431 using the L_1-norm to detect KDE changes. The sensitivity of the L_1-norm is lower than the one of the RMS bands.

The actual number of anomalies that is returned by the method is for many parameters still too high to be practical on an individual basis. As for the case of the RMS bands, however, we can use the anomaly detection in a statistical sense. That is, we let the software issue an alert when a significant number of parameters flag an anomaly simultaneously.

5.5 KDE changes summary

This Section tested algorithms that would enable to detect significant and unexpected changes in the KDE over orbital time. These algorithms are thought to be complementary to the single outlier detection algorithm presented in Section 4. We tested three different methods:

1. Confidence bands
2. RMS bands
3. Area variation (L_1-norm)

For the algorithm using confidence bands we found that it was extremely sensitive, especially for the well-sampled telemetry parameters, and that it therefore should only be used for those parameters that ought to be very stable. Comparing the algorithms that use RMS bands or area variations, we did not find a clear winner. Both clearly show potential. They tend to be rather sensitive, but when focusing on the more “quiet” telemetry parameters, and using them in a statistical (that is: combined) sense, we were able to detect the beginning and the end of the orbital manoeuvres without any prior knowledge.

6 Novelty detection based on Poincaré Plots

This Section describes our investigations around novelty detection using Poincaré plots (hereafter PP). Given a time series $\{y_i\}$ sampled at points $i=0,\ldots,N-1$, the PP will display y_i on the abscissa, and y_{i+L} on the ordinate. Hence the PP is a 2D plot taking the current values of the time series on the x-axis, and the future values of it on the y-axis. It is above all a simple way to represent the transitions in a time series graphically, in a way that all identical transitions in the signal collapse to a single point.

Displaying the data in a PP sometimes dramatically reduces the apparent complexity of the original signal to just a few data points. This is mostly true a priori in quantified time-series, especially in status words or 'enumerated' time-series (also called categorised datasets in statistical publications).

In addition to reducing the signal complexity, the Poincaré plot also gives access to rather intricate signal properties at a glance. Figure 9 shows a time series where the signal only takes 6 values. The PP shows that the values are grouped in two sets: A=[0,1,2] and B=[3,4,5], with an interesting transition pattern between the groups. First, transitions from A to B only occur from 0. No transition starting from 1 or 2 ever reaches B. On the way down, the only "possible" transition is 5 → 2. No other route from B to A is ever taken. This information that the PP just offers in a glance would be close to impossible to directly retrieve from the original time series.

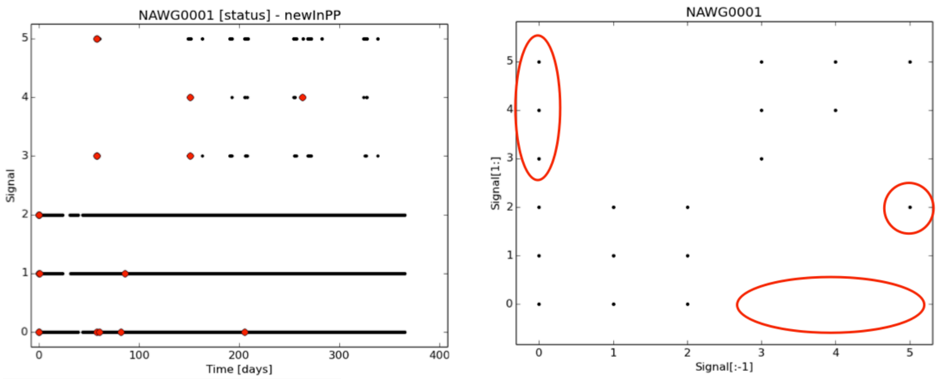

Figure 9: NAWG0001. Left: time series. The red dots identify the first time any given transition occurs. Right: Poincaré plot. The upper levels are exclusively reached starting from value 0. From high to low values, the only existing transition is 5 → 2.

The analysis of the Poincaré plots is useful to detect novelty in enumerated signals. In a PP, a stable signal is represented by a single point, on the first bisector (y=x). A signal alternating between two values a & b results in a PP with only two points, representing the transitions [a,b] and [b,a]. Hence, in simple situations, one will expect to find mostly collections of points on the first bisector, or symmetric around it. Now think of a PP building up with time, in a growing window. After a while, most operating modes of the S/C have been visited, and no new transitions are expected. Figure 10 displays the feature 'is new in the Poincaré plot' on a Venus Express parameter time series. One can indeed see that most transitions are visited early in the mission, but one novelty appears after ~300 days, linked to an unusual signal pattern, skipping one intermediate state during a transition, that had always been visited until then.

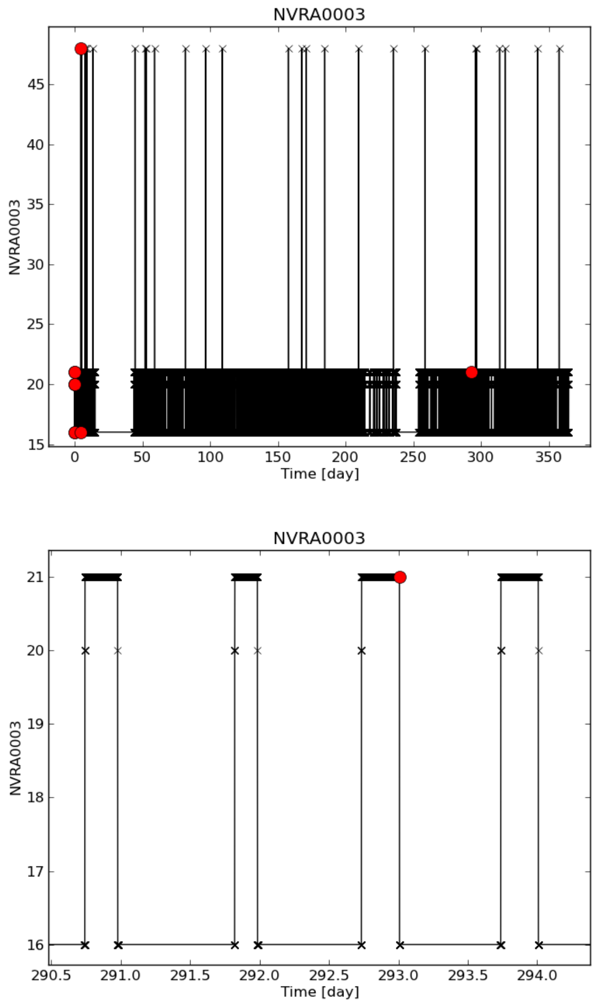

Figure 10: Novelty detection feature "*is new in Poincaré plot*", computed for lag=1 in NVRA0003. The top panel shows the clustering of the 'detections' at the very start of the time series. The bottom panel shows how the 'standard pattern' is altered after 293 days, when one state is skipped on the downward transition.

Can PPs be useful for all time series? In general terms, one would not expect the PP of a continuous signal to display peculiar features. There is nevertheless a wealth of information to be retrieved from the way it changes. Figure 11 shows Venus Express parameter NACW0G05, and its PP of lag 1. The first million samples (70 days) have been highlighted in red. Whereas this period appears absolutely nominal at first glance from the TS itself, it is clear from its PP that some transitions occur then, which never reappear.

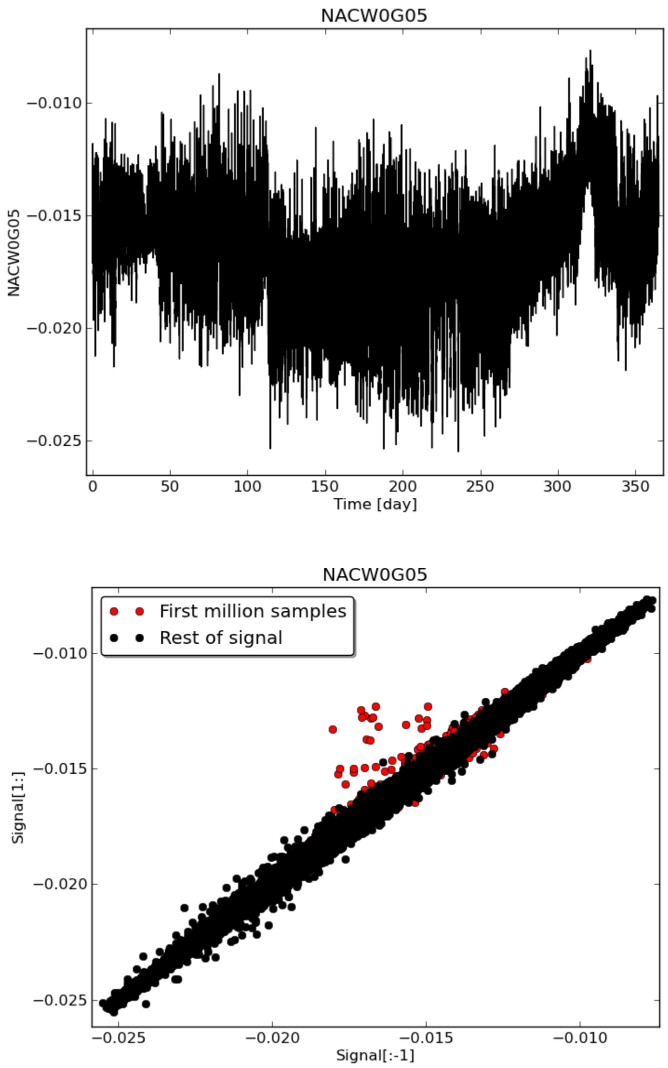

Figure 11: top: NACW0G05 vs time [sec]. Bottom: Poincaré plot of lag=1. The red dots correspond to the first million samples (~70 days, most of them hidden), and contain all the outliers.

The approach to analyse the PP of enumerated (incl. status) and 'continuous' parameters is different. Typically the methods we investigated for enumerated parameters do not make much sense as soon as the number of possible transitions increases beyond a few tens

If we define an enumerated parameter as one visiting maximum 8 possible values, it means it has a maximum of 56 possible transitions between these values. In the Venus Express mission history we find that 545 of the 1666 non-constant parameters obey this criterion (max 56 different transitions observed), i.e. about 1/3.

The distribution of the numbers of transitions in status and non-status parameters are compared in Figure 12.

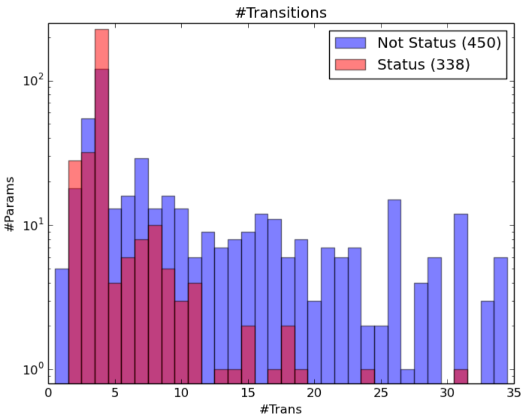

Figure 12: Distribution of the number of parameters with less than 35 distinct transitions. Few status parameters display more than 10 transitions, while the distribution is flatter for non-status parameters.

Features based on the Poincaré Plot display a lot of potential for novelty detection. This is especially true for enumerated/categorised parameters, where the signal complexity is dramatically reduced in the PP, and can be analysed with simple features like 'isNewInPoincaré'. We consider this to be the case not only for status words, but also for any parameter displaying a limited set of transitions, i.e. of points in the PP. We estimate that about 1/3 of the non-status parameters fall in this category.

For continuous signal, a large variety of more complex secondary features can be envisaged. We have explored the potential of indices based on the most occurring transitions, but the Poincaré Plot bears much more potential, on which the details are still to be finalised (e.g. structural analysis of outliers) or to be explored (e.g. detection and analysis of pseudo-frequencies).

7 Command effect analysis

This Section looks into learning correlations between commands and changes in telemetry values or their features. For all command executions found in the derived parameters from the auxiliary data, the typical 'footprint' of the command is determined by finding the parameters and parameter features that consistently show a similar before/after-command-change. This can be applied in an operational environment to learn the nominal effect of a command as the mission (or the AIV phase) evolves. This opens possibilities for an operational setup where the before/after-command-change of all parameters/features is compared to the nominal command effect and can help the spacecraft operator to be warned about potential problems with uplinked commands or command sequences.

Our approach is to study per command whether and how often we detect a significant change in a raw parameter and/or proposed features after the command was executed with respect to before. We compare the parameter over a window of 1 hour around the command. This indicates how likely it is for a raw parameter or feature to change (or not to change) when a command is executed, which can be used to predict the behaviour of parameters when the same command will be sent in the future.

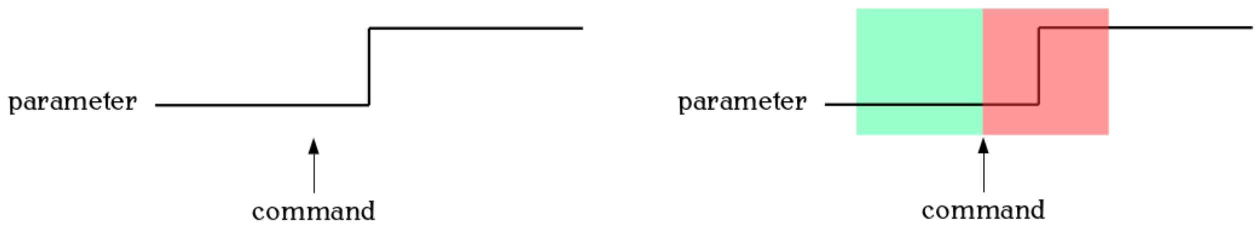

Figure 13: Change in raw parameter or feature value in response to the execution of a command. On the right-hand side the compared time windows are marked.

To define what a significant change in raw parameter or feature value is, we make a random selection of 20 such changes over the whole time series and calculate the standard deviation **σ** over these differences. Any change associated with a command greater than 1.5 times **σ** is considered significant.

We compare the before/after command difference of a parameter in the raw signal, and in three features: skewness, mean subtracted parameter and median subtracted parameter. The rationale for this choice is to have a robust and easy to compute feature describing the data distribution (skewness) and a detrended feature (mean subtracted, median subtracted).

A parameter is said to be affected by a given command if a significant change in raw time series and/or in at least one of the proposed features is observed in at least 95% of the times that command is executed.

The histogram in Figure 14 shows how many parameters are typically affected in the raw values by the different telecommands in the Venus Express mission history. The features provide additional sensitivity. This is illustrated in Figure 15, where the effect of a telecommand is not visible in the raw parameter value, but is clearly visible in the skewness feature.

.

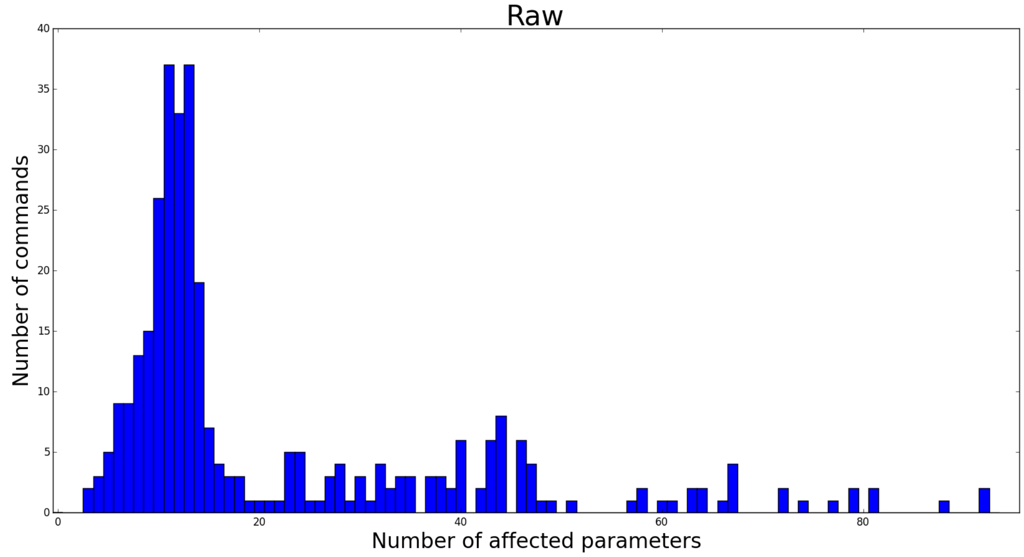

Figure 14: Histogram showing how many parameters the commands affect in the raw time series.

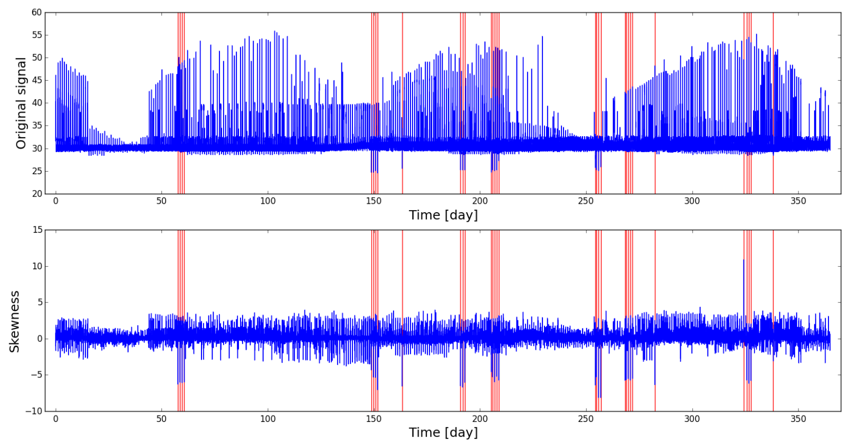

Figure 15: The raw parameter (top) and skewness (bottom) time series of parameter NAAD0401 (in blue). Over-plotted in red are the execution times of command ZAC02092. The red bands (rather than vertical lines) indicate the command is executed multiple times within a short timespan. In the raw parameter this commands results in a significant change in only 28.1% of the cases, as opposed to 96.6% in the skewness feature. Only the last execution of the command does not result in a significant change in skewness.

8 Lessons learned and possibilities for follow-up studies

We developed different techniques for novelty detection in spacecraft telemetry parameters and derived parameters. The Kernel Density Estimate (KDE) of data provides an accurate and robust description of the data distribution that can be used for novelty detection via single outlier detection. The KDE also provides additional novelty detection possibilities in terms of changing data distribution and recognition of epochs of characteristic data distribution. Future work based on this study can include refining the classification of parameters to automatically choose appropriately sensitive features, experiments with an adaptive anomaly threshold, influence of gaps in the time series, and anomaly detection of absent values in the distribution.

The representation of time series in Poincaré plots and the automatic analysis techniques to interpret changes in the Poincaré plots provide a complementary power for novelty detection, not only in continuous data but particularly for enumerated or status time series. Further studies on Poincaré plots could elaborate on transition frequency analysis (as opposed to baseline analysis), pseudo frequencies and filling-factor outlier detections.

Our work on Command Effect Analysis shows potential as a powerful tool to help spacecraft operators to assess the correct execution of commands with fairly simple (hence realistic to implement in an operational environment) algorithms. It could also be of use to complement the manually edited spacecraft in-flight operational procedures, which typically include minimal fail/pass criteria to assess the correct execution of a command. Follow-up studies could investigate how to weigh the importance of a parameter/feature in the effect of a command, how to refine the resolution of a command based on the command parameter values, how to use information on correlated commands, how to use additional features that would be more sensitive to a flag going up for a limited time, then go down again; features that are more sensitive to a gradient change, etc. and investigate different time scales to assess the before/after effect and assess complementarity of these time scales

In general, the study focussed around characterising parameter time series via the statistical distribution of data values. There is additional potential to characterise time series in the time domain, analysing the 'tune' (local shape) of the signal rather than its spectrum and to develop 'regularity' features, including various measures of entropy, fractal index of the signal etc.

A. Acronyms

AIV	Assembly, Integration and Verification
ASSC	Automated Spacecraft Status Characterisation
CSV	Comma Separated Values
ESOC	European Space Operations Centre
LoOP	Local Outlier Probability
MCN	Minimum Cell Number
MIC	Maximal Information Coefficient
MINE	Maximal Information-based Non-parametric Exploration
MST	Minimum Spanning Tree
MUST	Mission Utility and Support Tools
NaN	Not A Number
PP	Poincaré Plot
SD1	Standard deviation around the first bisector
SLURM	Simple Linux Utility for Resource Management
TC	TeleCommand
TS	Time Series
VEX	Venus Express