DRMUST: AUTOMATING THE ANOMALY INVESTIGATION FIRST-CUT

José-Antonio Martínez Heras(1), Kar Lam Yeung(2), Alessandro Donati(2), Bruno Sousa(2), Norbert Keil(3)

(1) Black Hat S.L., Santa María de Gracia 18, 14002 Córdoba, Spain, Email: j.martinez@blackhat.es
(2) ESA/ESOC – OPS-HSC, Robert Bosch Strasse 5, 64293 Darmstadt, Germany, Email: {kar.lam.yeung, alessandro.donati, bruno.sousa}@esa.int
(3) LSE Space AG, Münchner Str. 20, 82234 Wessling, Germany, Email: norbert.keil@esa.int

ABSTRACT

Whenever a flight control team is faced with an anomaly they need to understand what its effects are and try to identify its cause. Understanding the effects may lead to actions to minimize them; understanding the cause allows in some cases to avoid it from happening again in the future.

The anomaly investigation current approach is based on the knowledge, experience and intuition from engineers. They hypothesise which parameters could be involved in the anomaly and then perform the analysis to prove or discard their hypothesis.

DrMUST's approach consists of automatically finding the involved parameters for the engineer, even those that had not yet been considered. It uses speech recognition techniques to identify nominal anomalies and a novel approach to perform correlation analysis based on statistical features. DrMUST can be used not only to support anomaly investigation but also in performing characterizations, model tuning and finding similar patterns in telemetry.

This paper describes what DrMUST is, how it works and its evaluation with Venus Express in three different scenarios. Finally, the flight control team provides its operational assessment.

1. INTRODUCTION

Anomaly investigation is part of the routine diagnostic task of flight control engineers. When an anomaly is detected (e.g. when a particular parameter crosses a noticeable threshold), it is common practice to search along the telemetry history for similar behavioural patterns in order to characterize the anomaly. By analyzing the time periods when the anomaly happened in the past, we may be able to identify its causes and eventually prevent it from happening again in the future.

The anomaly investigation process can be very labour intensive: many different parameters need to be analysed to identify possible correlations with the observed anomaly. DrMUST is an ESA's Advanced Mission Concepts and Technologies Office initiative, based on an idea and need presented by the Venus Express Flight Control team, to support ESA operations engineers in general by reducing the time and resources needed for anomaly investigation.

Although DrMUST has been designed with the goal of supporting anomaly investigation; it can also be used to perform system or subsystem characterization. This process helps engineers in identifying potential areas of concern when operating the spacecraft in different modes.

The rest of the paper is organized as follows: section 2 provides with a literature review, section 3 states what DrMUST does and how it is intended to be used, section 4 describes how DrMUST work, section 5 discusses implementations details, section 6 is devoted to evaluate DrMUST in three different cases using Venus Express, section 7 consists on Venus Express Flight Control Team operational assessment, section 8 concludes with ideas for future DrMUST extensions.

2. BACKGROUND AND RELATED WORK

Upon development of the idea for this new project for aiding anomaly analysis, we realised that two components would be required: the ability to find similar events that had happened in the past and the ability to find which parameters may be involved in the anomaly.

For the case of finding similar events, it is common to compare sequences with each other using distance measures such as Euclidean distance to determine similarity. These approaches use a sliding window to index the data and then compute the similarity of the created sequences against the original to find the most similar or dissimilar [6]. E. Keogh et al. [7] has developed at technique called SAX that makes this search particularly efficient.

However, the pattern matching algorithm that we needed could not be based on Euclidean distance as the patterns are seldom identical. In the telemetry data we are analysing, we should allow for small misalignments

Proc. 'IJCAI–09 Workshop on Artificial Intelligence in Space', Pasadena, California, US 17–18 July 2009 (ESA SP-673, September 2009)

in time and value. The same problem is solved by the speech recognition community [3, 4] by using a technique called Dynamic Time Warping (DTW). E. Keogh [5] provides an efficient way to perform DTW matches and also provides with an exact indexing approach. In this work, DTW is described as a more robust approach to measuring distance (difference) between time series as it allows for sequences that are not aligned in time to be considered as similar. This is the distance measure we use to rank sub-sequences in the data as being similar to a given pattern.

While there are other approaches matching our work more exactly [8], they did not meet our needs as they required increased complexity, for example producing an index of the time series before search [9]. As we expect to have different time periods lengths and parameters for each query, creating the index would not save time in this case.

There is also widely published work for anomaly detection in a number of different applications, for example in [6]. The approaches that are usually used for anomaly detection in a time series assumes that you are looking for the most unusual sequence to appear over a period of time. The user defined anomalies we would be working with do not necessarily conform to this definition so we had to return the anomaly matches in a different way.

Regarding the cases of finding the parameters that may be involved in the anomaly, we did not find in the literature an approach that would provide a simple yet systematic and efficient mechanism to suggest these correlations that would be robust enough to work for the many different parameters involved.

3. WHAT IS DRMUST?

DrMUST is a set of techniques implemented in a computer prototype that allows engineers to efficiently:

- find similar patterns in a given telemetry parameter from a large time period (in the order of years)
- find the telemetry parameters that are involved in a relevant time period (e.g. anomaly) from a large number of parameters (in the order of thousands)

The work-flow is the following:

1. The user defines a time period for a given parameter that signals a behaviour to be investigated, e.g. an anomaly period (A in Fig. 1)
2. DrMUST finds other occurrences of the same parameter with a similar pattern (B in Fig. 1).
3. DrMUST performs correlation analysis by finding parameters that behave consistently in the anomaly periods (defined by A and B) and differently in nominal periods.
4. DrMUST provides the user with the list of parameters that are potentially involved in the anomaly.
5. The user evaluates the results, discarding irrelevant correlations (e.g. coincidental) and classifying the remaining correlations (e.g. cause, effect, knock-on effect).

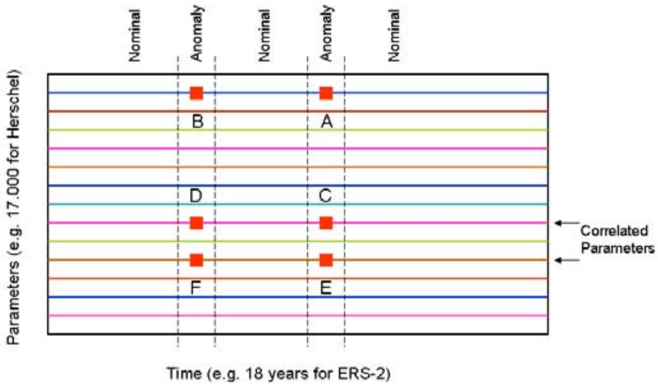

Figure 1. DrMUST can find similar anomalies for a given parameter behaviour and find other parameters involved in the anomaly

There are occasions where users can go directly to step 3 of this work flow:

- When investigating a known recurrent anomaly. In this case the anomaly and nominal periods are known.
- When performing characterization. Users know the time periods they want to characterize (used as anomaly periods) in contrast to other periods (used as nominal periods).

DrMUST (pronounced 'Doctor MUST') is named after MUST, the Mission Utilities & Support Tools. MUST [1, 2] is a collection of tools that support the analysis, visualization, exploration and exportation of telemetry and ancillary mission data. DrMUST uses MUST as data provider.

4. HOW DRMUST WORKS?

4.1. Finding similar periods

Anomalies are commonly referred to as deviation from expected behaviour. They are usually recognised by the Flight Control Team because the behaviour of certain parameter(s) is not one that is expected. For example, the anomaly is found from an out-of-limits check or from recognizing a different shape in the parameter time series.

We can classify anomalies as first-time anomalies or recurrent anomalies. We refer to recurrent anomalies as those where we identified similar occurrences in the past and first-time anomalies as those without previous occurrences. In some cases, however, first-time anomalous behaviours did indeed happen more times in the past. They can pass unnoticed due to a number of reasons, such as not exceeding any out-of-limits.

DrMUST provides the capability of searching for patterns (time series behaviour characterised by a shape) in a given parameter over a large amount of time. It allows us to answer the questions: "Is it really the first time that this anomaly happened since the beginning of mission?", "Do we know all the occurrences of a recurrent anomaly?” It finds all time periods for a given parameter where the pattern is similar to the one specified.

It is difficult to find similar behaviours in a given time series since periods that engineers would recognize as having the same shape, in reality have misalignments in time and can have different scale. In order to cope with the uncertain nature of similar shapes, DrMUST uses Dynamic Time Wrapping (DTW) [3, 4]. DTW allows recognising of shapes as similar even with minor fluctuations such as misalignment, differing speeds or differing amplitude. DTW is computationally intensive; however we make use of recently developed optimizations [5] to make it efficient.

4.2. Correlations

DrMUST makes two assumptions in order to find correlations among housekeeping telemetry parameters:

1. Parameters related to the anomaly behave similarly in all same-anomaly periods.
2. Parameters related to the anomaly will behave differently during anomaly and nominal periods.

DrMUST's solving approach consists of scanning every parameter and suggesting to the users those parameters with a similar behaviour during anomaly periods and different behaviour during nominal periods.

Similar behaviours in anomaly periods are characterised by features. We consider that a parameter behaves similarly in all anomaly periods if at least one feature shows similar values (a small deviation). Different behaviours between anomaly and nominal periods are characterized by different feature values (a large deviation) for at least one of the features that showed consistent behaviour during the anomaly periods.

The features currently implemented in DrMUST are: average, standard deviation, maximum, minimum, range, slope, maximum minus slope, minimum minus slope, skewness, kurtosis. DrMUST is designed in a modular way that easily allows for the addition of more features when necessary. We found this approach more useful than classical regression or finding similarity with DTW since many different shapes can characterize both anomalous and nominal behaviours. Requiring that at least one feature is common in anomaly periods and different in nominal periods is a more robust approach that works in more cases when faced with real telemetry data.

In order to find correlations, the user must define at least 2 anomaly periods and 1 nominal period. Defining more time periods of each kind will help in reducing the number of coincidental correlations. The two or more anomaly periods could already be known (e.g. due to a recurrent anomaly) or they can be found by using the pattern matching as described for finding similar shapes. If the anomaly is truly new (we can confirm that no similar behaviour happened in the past), the only available option is to split the anomaly period in 2 time periods in order to meet the 2 anomaly periods requirement. This will increase the chances of getting coincidental correlations but, at least, we will be able to find some correlations.

5. IMPLEMENTATION

5.1. Similarity search

We will not go into the details of DTW but instead refer to the research [5] for full details of the algorithm, and describe how it is a useful measure of similarity for parameter data. We found DTW to be more suitable than using Euclidean distance as it allows for the discovery of sequences that have an overall similar shape but are not necessarily aligned in time. With spacecraft data there tends to be small fluctuations present and these can lead to a failed pattern match if it is not taken into account. DTW allows shapes that are similar but out of phase to be matched correctly. Lower bounding was very important to speed-up the pattern matching process (computing DTW distances is much more computationally expensive than computing Euclidean distances). Lower bounding is the use of an approximation function that is less computationally intensive than the full distance calculation that returns an approximate result, guaranteed to be less than or equal to the real value. We used the LB_Keogh function as in [10] as it has been proven to guarantee no false dismissals and returns results close to the real distance function.

5.1.1. Pre-processing

There are a few steps that are required when working with spacecraft parameter data in order for time series search to work properly. A parameter can have a varying sampling rate over time, meaning that a direct comparison between one subsequence and another are not possible unless they are first resampled at the same rate. The resampling is done by linear interpolation to convert to the required number of samples. DrMUST uses the number of samples of the original query. During testing it was found that gaps in the data could cause incorrect matches. These are dealt with by skipping over data if there are no values recorded over a specified amount of time.

5.1.2. Range and scale cut-off

DTW finds shape matches that are independent of range and scale. Often, an engineer is looking for a behavior of a certain magnitude (e.g. spikes). Values and spikes of the same shape that are outside of these values are not useful. That is why we implemented the option of excluding matches that are outside of a user specified minimum or maximum range, minimum or maximum scale, or any combination of these. The parameter data tends not to be completely flat when there is little to no change but has fluctuations of a small magnitude. These fluctuations can often match the shape of a sequence while being meaningless to the pattern search. Normally a minimum range is required to eliminate these false matches. Maximum range is sometimes required to narrow down the true anomalies. By checking if the subsequence range is acceptable first, and discarding the ones that not meet the criteria, the number of required DTW computations is reduced further thus speeding up computation.

5.1.3. Best k matches

As DTW distance values are relative to the problem, there is no immediate way of identifying whether the match is valid or not (e.g. we cannot define an unique threshold distance that is valid for all problems). There may not be any similar shapes existing in the data or the matches returned could be invalid in another way. We offered the option of number of matches to be returned as a user input and rank the matches by distance and range, if range has been selected. Then the engineer can decide how many of the returned results can be used for further analysis.

So far we have found this approach to be adequate for the time series used and very good matches were returned. Since the matches could be due to seasonal variation or intended manoeuvres, we cannot discount matches from being nominal behaviour so they need to be confirmed by an engineer.

5.2. Correlation

The user needs to specify at least 2 time periods where the same anomaly occurred and at least 1 nominal period. Best results, in terms of reduced coincidental correlations, are obtained when the nominal periods are selected to be close to the anomaly periods. Increasing the number of time periods, both nominal and anomaly, generally improves the relevance of the parameters found in the correlation process.

The anomalous time periods can be specified either manually (e.g. known recurrent anomaly or when performing characterization) or using the results obtained from the similarity search. The nominal time periods are specified manually.

DrMUST then scans all telemetry parameters available in MUST, around 7000 for Venus Express, for correlations. DrMUST requires a similar behaviour during anomalies that is different during nominal periods.

Currently the correlating stage can take over an hour as it is limited to the rate that data can be read over the network. However, such tasks can take months if done manually. The search can be done in the background unattended with the results returned when processing is complete. Also, locally stored data would not be subject to such a restriction and is instead limited by the speed of fetching the values from disk.

6. EVALUATION

DrMUST has been evaluated by Venus Express control engineers with three cases.

6.1. Characterization of solar arrays when working at their maximum power

The Venus Express Flight Control Team wanted to find occasions when the solar arrays were operating at their maximum power to estimate their efficiency. This does not occur very often because they are either fully illuminated and then produce excess power or they are sharply put in the shadow and produce no power at all. The goal was to find periods of a few minutes during which they had such an inclination and power demand that forced them to produce their maximum capacity and find which other parameters are affected in this scenario.

6.1.1. Finding similar shapes

The indication that the solar arrays are operating at their maximum power is given by a control signal provided by the power regulator, when its value lies between 7.5 and 10 V. An example time period showing the desired mode transition (to around 9V) over a period of four minutes was used to begin the search. DrMUST searched for 4-minute patterns in a time span covering 3 years. It successfully found 2 other time periods with similar behaviour to the given period. This process took less than 2 minutes including retrieval time.

Fig. 2 shows the original values used to start the search, Fig. 3 and 4 show the two valid matches that were found. Matches that have the same overall shape can be found and are not dismissed due to variations.

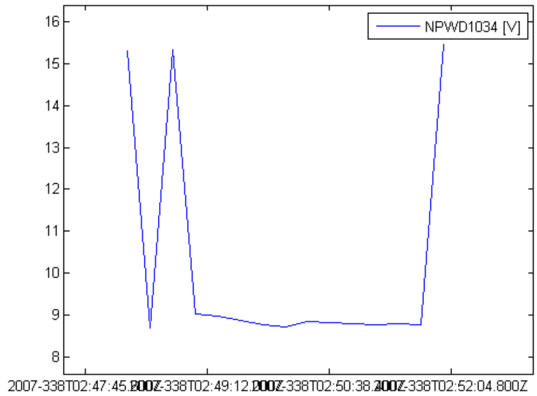

Figure 2. Shape search target

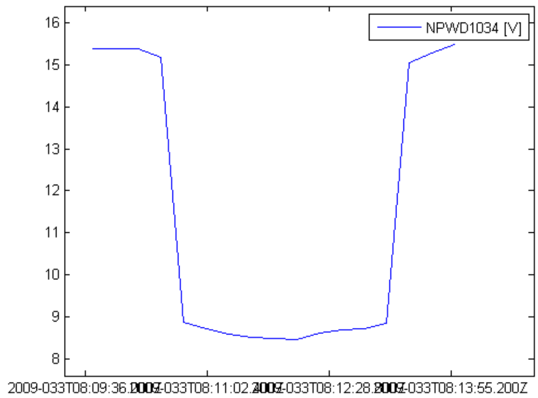

Figure 3. Shape first match

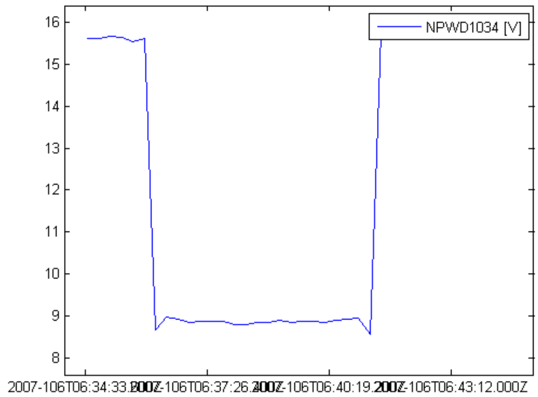

Figure 4. Shape second match

6.1.2. Finding correlated parameters

These 3 periods were used as "anomaly periods" and 4 other normal periods were used as "nominal periods". DrMUST correctly found the correlations expected by the Flight Control Team.

Fig. 5-8 show examples of user generated graphs to show the correlation. These examples are taken from the parameters found by DrMUST and show the types of features that can be considered in the search. Each graph shows three different time periods within the target or nominal ranges for the four minute periods.

Parameter NPWD1124 (Battery Discharge Current) was found to have features values that are similar to each other during the target time periods and different from nominal time periods for the features: average, maximum, range, standard deviation and maximum minus slope (Fig 5, 6). The reason why this parameter correlates is that whenever the solar arrays are operating at their maximum power, the excess power demand has to be provided by the batteries. So a small discharge ensues.

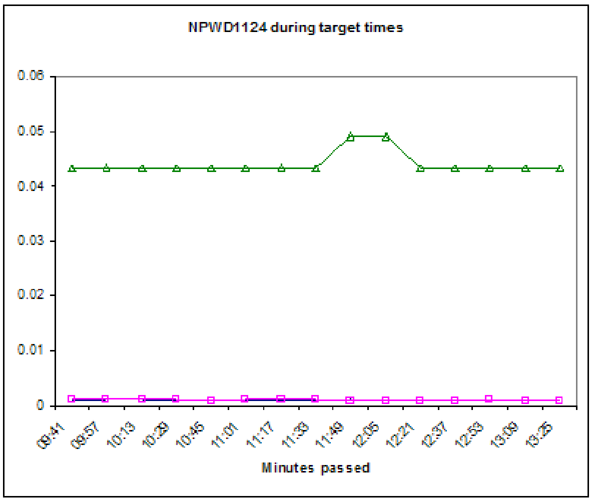

Figure 5. NPWD1124 target times.
The scale is close to 0

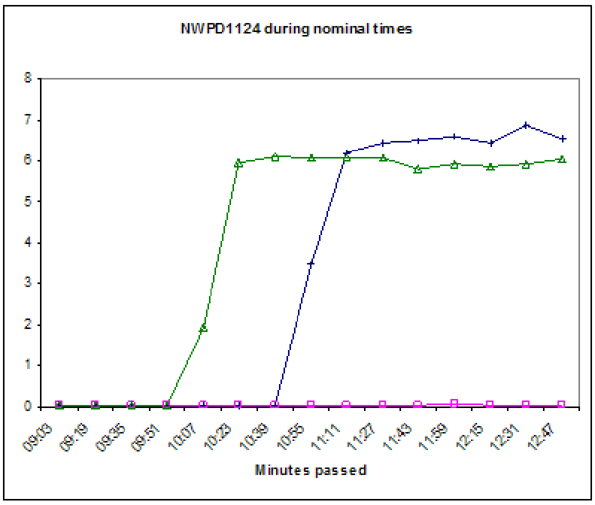

Figure 6. NPWD1124 nominal times

For parameter NPWD1294 (Battery Charge Current) the features that did not match in target and nominal times here were average, maximum and minimum minus slope (Fig 7, 8). Again this parameter correlates because after a discharge, the re-charge process is activated.

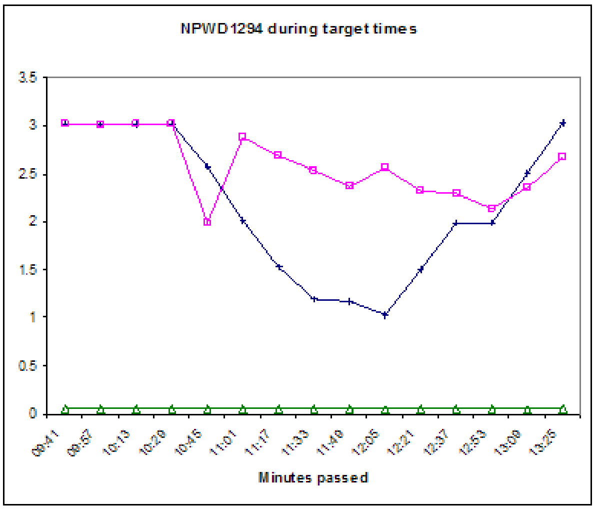

Figure 7. NPWD1294 target times

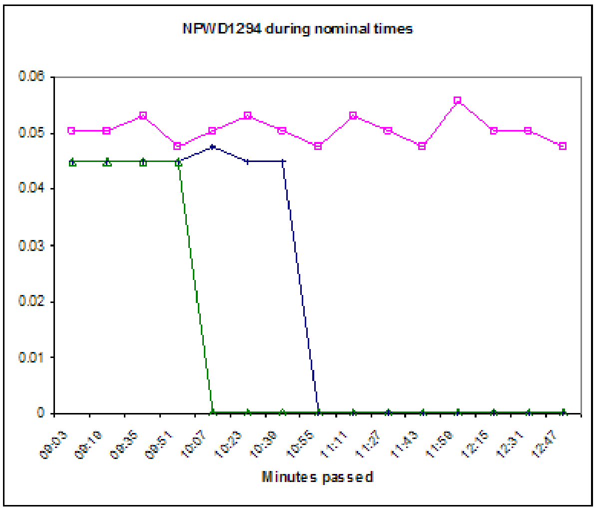

Figure 8. NPWD1294 nominal times
The scale is close to 0

6.2. Quadrature Characterisation

There is a phase in Venus Express call Quadrature when the angle Earth-Venus-Sun is within 80 and 110 degrees. In this phase, the Flight Control Engineers offset the Venus Express attitude during steady earth pointing periods to avoid direct illumination on the optics of a payload. This has knock-on effects on other aspects of the spacecraft, such as thermal environment, power management, wheel momentum management, etc.

DrMUST was used to characterise the overall behaviour of the spacecraft in terms of correlated parameters. These correlations are needed to detect any potential areas of concern. The results obtained by DrMUST matched the expectation of the Flight Control Team in terms of correlated parameters.

6.3. Anomaly investigation: reaction wheel 2 friction peaks

The anomaly used as a test case was the Reaction Wheel 2 (RW2) friction peaks on the Venus Express spacecraft. RW2 sometimes shows unexpected peaks in its friction. Here we were interested in determining when these peaks occur and what other parameters have correlated behaviours that could help identify causes or consequences of these friction spikes.

We were given a set of examples where this friction peak had occurred since January 2009. The peaks were known to be more common in certain months of early 2009. Using these periods we were able to identify that these peaks had occurred many times in the past, in addition to the investigated cases from 2009.

With these time periods as an input, no parameters were found that would cause the friction wheel spikes directly or indirectly and a domain expert confirmed their belief that the anomaly is likely to be internal to the wheel itself. While the discovered parameters showed no cause, some of the correlations that were found turned out to be an affect of the anomaly and showed that the friction problem impacts on pointing accuracy, an issue also proven by Flight Dynamics. Finding such issues before they cause further problems is advantageous and can aid with decision making for new operations strategies.

Fig. 9 shows two parameters that were found to correlate with the reaction wheel spikes. The green line shows the data for the reaction wheel, it can be seen that when the other two parameters show a flat line, the spikes are seen to occur. The values are normalised in order to show the effects at a reasonable scale.

During the same investigation, a model was devised to estimate the friction as a function of the commanded torque and temperature to be able to predict future trends. The model was able to follow the estimated friction on board (obtained in telemetry) except for some particular periods where it would deviate considerably. By running DrMUST over those periods it was possible to identify the source of the deviation. It resulted from an additional torque on the wheel introduced by the rotation of an electric motor that moves the shutter of a payload.

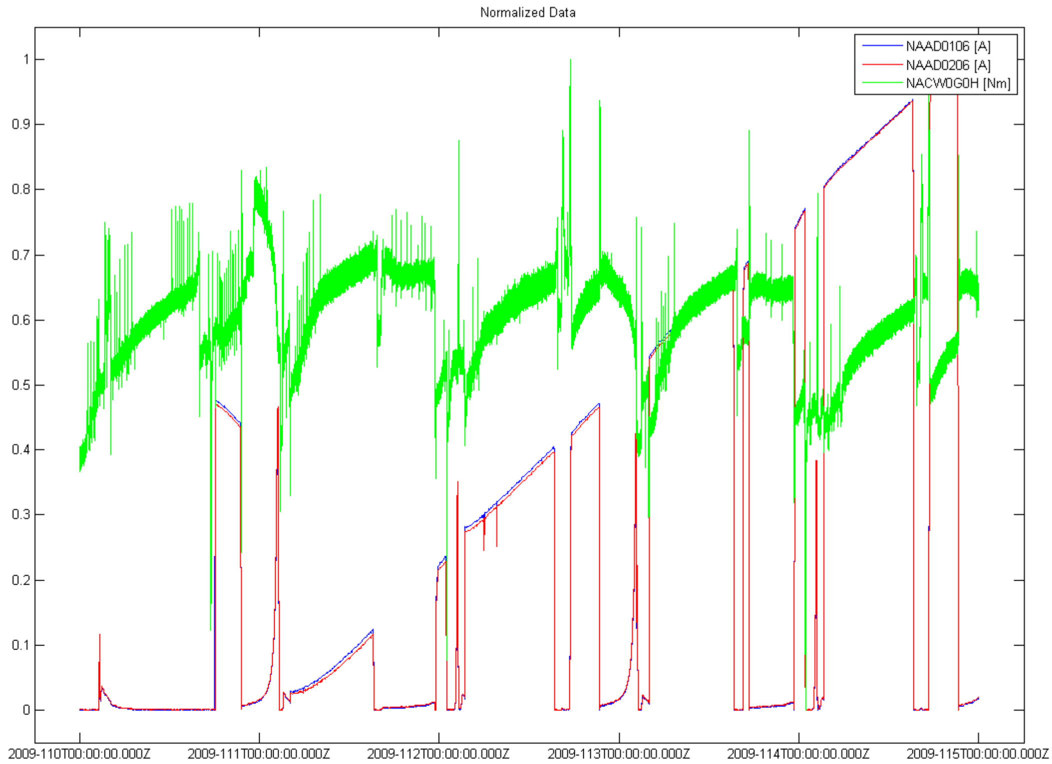

Figure 9, Reaction Wheel Correlations

7. OPERATIONAL ASSESSMENT

DrMUST assists in data analysis tasks which are very labour intensive. Before, it was required that the engineer would guess or hypothesise the parameters that could have a credible correlation to a specific behaviour and would then perform the analysis to prove or discard the correlation. DrMUST finds these correlations for the engineer, even those that had not yet been considered. It also helps eliminating hypotheses by showing a-priori that there is no significant correlation. The Venus Express engineers estimate that DrMUST can save up to 20% - 30% of the work required to perform anomaly investigation or phase characterization.

The Venus Express Flight Control Team plans to use DrMUST to perform anomaly investigations and characterizations of the overall spacecraft behavior over different phases. It is expected that DrMUST will be used more and more as Venus Express gets older since more anomalies are likely to occur. In addition, DrMUST will be a valuable tool during the Venus Express aero-breaking phase. It will assist engineers in characterizing new behaviors, supporting decision making in the adoption of new strategies and identifying potential hazardous situations.

Venus Express plan to incorporate DrMUST in support of the following activities: mission review boards, decision making concerning new operations strategies, anomaly handling, procedure modification, spacecraft or subsystem configuration, payload usage, subsystem deterioration management, behavioural model calibration.

8. FUTURE WORK

DrMUST is an on-going project with the following major prototyping activities in the pipeline:

- Inclusion of asynchronous data: currently, all the correlations are performed only on housekeeping telemetry parameters that are available with certain continuity. The next step is to include also asynchronous data such as spacecraft events and telecommands. This is particularly interesting to detect anomalies due to certain command or to learn, for instance, that a certain anomaly is anticipated by an event.
- Correlations classification: correlations can be classified as cause, effects, knock-on effects and coincidental. However, DrMUST cannot tell which one is what. We are working out a new concept that will provide users with an initial indication of what kind of correlation it might be.
- Finding the parameters involved in an anomaly could be the first step for a more elaborate analysis using AI techniques. For instance, the parameters suggested by DrMUST could be the starting point for a clustering based monitoring system.

9. REFERENCES

1. Baumgartner A., Martinez-Heras J. A., Donati A. & Quintana-Claramunt M., "MUST - A Platform for Introducing Innovative Technologies in Operations", *iSAIRAS 2005*, Munich, Germany
2. Martinez-Heras J. A., Baumgartner A., Donati A., "MUST: Mission Utility and Support Tools", *Proceedings DASIA 2005 Conference*, Edinburgh, UK
3. Itakura F. (1975) Minimum prediction residual principle applied to speech recognition. *IEEE Trans Acoustics Speech Signal Process ASSP* **23**:52–72
4. Sakoe H. & Chiba S. (1978) Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans Acoustics Speech Signal Process ASSP* **26**:43–49
5. Keogh, E. (2002). Exact indexing of dynamic time warping. In *28th International Conference on Very Large Data Bases*. Hong Kong. pp 406-417
6. Chiu, B., Keogh, E. & Lonardi, S.: Probabilistic discovery of time series motifs. In: *Proc. ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, Washington (2003)
7. Lin J., Keogh E., Wei L. & Lonardi S. Experiencing SAX: a novel symbolic representation of time series - *Data Mining and Knowledge Discovery*, 2007 - Springer
8. Faloutsos C., Ranganathan M., & Manolopoulos Y. Fast subsequence matching in time-series databases. *Proc. ACM SIGMOD*, pages 419--429, May 1994.
9. Shieh J. & Keogh E., iSAX: Indexing and Mining Terabyte Sized Time Series, *SIGKDD*. pp 623-631 2008.
10. Keogh E. J., Wei L., Xi X., Lee S.-H. & Vlachos M.: LB_Keogh Supports Exact Indexing of Shapes under Rotation Invariance with Arbitrary Representations and Distance Measures, *VLDB* 2006 pp. 882-893,