Predicting Venus Express Thermal Power Consumption

Hugo Penedones, Bruno Sousa, Alessandro Donati

ESOC/ESA European Space Operations Center, European Space Agency, Darmstadt, Germany

José Martínez-Heras

Black Hat S.L., Córdoba, Spain

Venus Express is a mission of the European Space Agency with the goal of studying Venus atmosphere. The spacecraft has a thermal system that keeps certain instruments within an operational temperature range. This system consumes some of the electrical power available, but it's hard to estimate how much. We propose a Data Mining and Machine Learning model that significantly improves the accuracy of the current predictions in the Mission Planning System. This should allow a safer planning of resource allocation and eventually an increase in science return.

I. Introduction

I.A. Venus Express mission

Venus Express (VEX) is the first Venus exploration mission of the European Space Agency (ESA). It was launched on November 9^{th} 2005 by a Soyuz-Fregat from Baikonur, Kazakhstan. After a journey of about five months it arrived to its final destination, where it keeps a 24-hour elliptical, quasi-polar orbit. At its closest, Venus Express reaches an altitude of 250 kilometres and at its furthest, it is 66000 kilometres away from the planet.

The spacecraft has been carrying global investigation of Venusian atmosphere in terms of its structure, composition and dynamics. To achieve that, it uses a set of spectrometers, spectro-imagers and imagers, covering a wavelength range from UV to thermal IR, along with a full plasma analyser.[2] VEX has a total of seven scientific instruments.

The Venus Express Mission Operations Centre (VMOC) is located at the European Space Operations Centre (ESOC) in Darmstadt, Germany. Communications with the spacecraft are mostly done using the ESA deep space ground station located in Cebreros, near Madrid, Spain.[1]

I.B. Problem Overview

The spacecraft uses electrical power coming from the solar arrays (or batteries, during eclipses) not only to supply the platform units, but also the thermal subsystem, which keeps the entire spacecraft within a temperature range. The remaining available power can be used by the payloads to do science operations.

The amount of power that will be required by the platform units can easily be determined. However, the power that the thermal subsystem will consume is dependent on the heaters, whose behavior depends both on external (e.g. sun distance, solar aspect angle) and internal heat sources (e.g. units' status, payloads' status). Therefore the consumption can only be estimated. This is done during the planning phase in order to reserve power for the thermal subsystem.

The goal of the project described in this paper was to forecast the power consumption profile of the thermal subsystem for a given time period in the future. An accurate power usage estimation model at planning level would reduce the current power margin allocated to the thermal subsystem, still complying with the safety margins for avoiding undesired safe mode reconfiguration due to excessive bus power demand. The gain in power will allow an increased level of science operations, and consequently of science return.

I.C. VEX Thermal sub-system

The purpose of the thermal subsystem is to maintain all the componets of the spacecraft within their respective operational temperatures. Venus Express uses both passive (e.g. paintings, insulation, etc.) and active (heater lines) mechanisms to fulfil this objective.

A *heater line* is an electrical component made of a heater and a thermostat (see Figure 1).

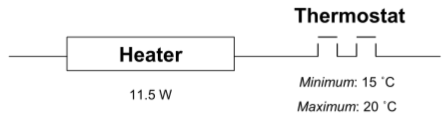

Figure 1: Example of heater line

Heaters are devices that can transform electrical power into heat. When they are active, they consume a fixed and known amount of power.

Thermostats are devices that control the activation or deactivation of heaters based on the temperature they measure. When the temperature at a thermostat goes below the minimum threshold, the associated heater is switched ON, whereas if the temperature goes above the maximum, the heater is switched OFF.

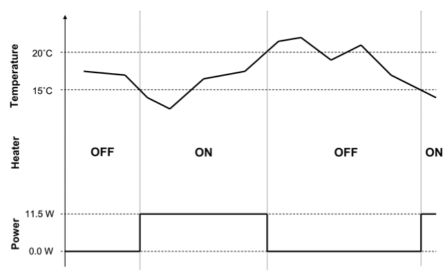

Figure 2: Thermostat behavior

The spacecraft has many heater lines powered by the Power Distribution Unit (PDU). This unit is responsible for delivering the power to spacecraft's units. Latching Current Limiters (LCL) are devices that control units that might be switched off. They are ON/OFF switchable current limiters for non-essential loads. Let's call LCLH to a LCL that controls a set of heater lines (see Figure 3)

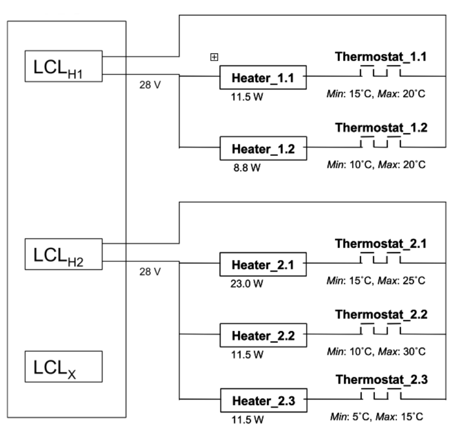

Figure 3: Thermal subsytem

Every heater line behaves independently based solely on the actions triggered by its thermostat. When the thermostat decides that its heater should be switched on, it closes the circuit and the heater demands power to the PDU through its LCLH.

I.D. Why is the problem hard?

Modeling the heater power consumption profile is hard because we are dealing with a complex, dynamic and recurrent system.

Note that the temperatures sensed by the thermostats depend on many factors: intensity of Sun radiation, Venus albedo effect, eclipses, spacecraft orientation, heat produced internally by the payloads, transmitters, batteries, etc. The thermal subsystem itself, will affect temperatures that command its own behavior, which creates a time-lagged feedback effect. Moreover, we don't have reliable predictions for some of these components (such as Venus albedo effect, or battery usage), which forbids us of including them as inputs to the model.

On top of that, analyzing telemetry data means that we have to deal with time series with millions of points, some data gaps, noise and different sampling periods.

I.E. Current Model

The current model in the Mission Planning System is a function with a single input attribute: the pointing type. The pointing type can be seen as an high level description of the operational state and orientation of the spacecraft in a given time period. For example, the spacecraft can be pointing Earth, pointing Nadir, performing a "slew", etc. (see Table 1 for complete list).

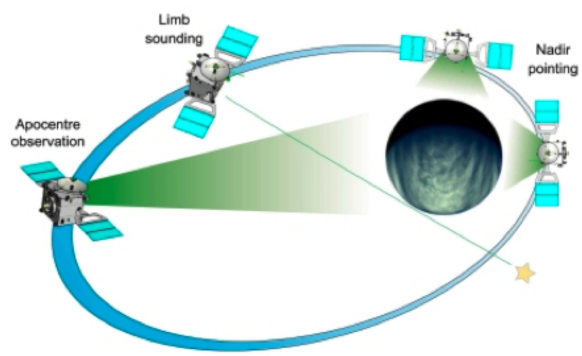

Figure 4: VEX Orbit and some pointing types

This simple forecast model was obtained doing physical experiments while Venus Express was being developed by ESA and the industry. It can be summarized in the following table:

Pointing Type	**Power (W)**
EARTH	146.2
SLEW	188.7
MAINT_OCM MAINT_WOL default	157.2
INERTIAL NADIR NADRI_POW MOSAIC CUSTOM ALONG_TRACK ACROSS_TRACK	156.8

Table 1. Power Estimation for each Pointing type

However, once the spacecraft is flying and sending telemetry, the over-simplification of the *a-priori* model becomes evident, as Figure 5 illustrates.

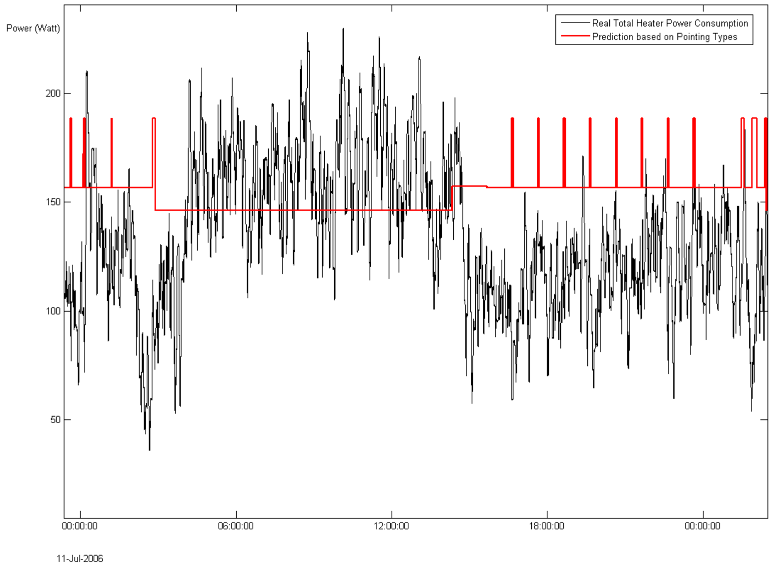

Figure 5: Real data (black) and current model prediction (red)

In fact, any predictive model based on pointing types only would be prone to fail, given that there is a big dispersion of heater power consumptions for each pointing type. The histogram of Figure 6 shows the wide range of heater power consumptions for the "pointing EARTH" situation, based on real telemetry data.

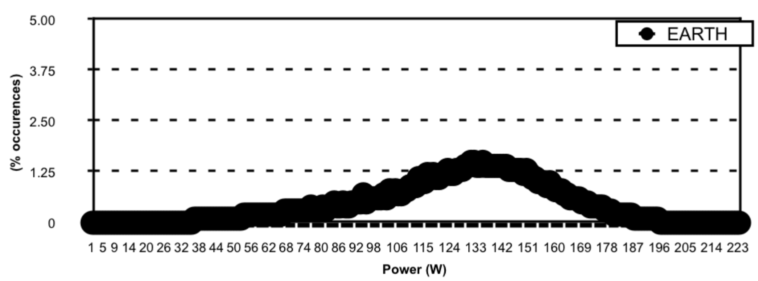

Figure 6: Histogram of heaters power consumption during pointing EARTH time blocks

Clearly, the prediction of 146.2W is not good enough, because when pointing Earth, the thermal subsystem might as well be using less than 70W or more than 200W and there is no way to distinguish those situations just based on pointing types.

It becomes now clear the need for *a-posteriori*, or *data-driven* models, that try to better match reality. We describe our approach to pursue this goal in the next sections.

II. A Data Mining Approach

Data Mining has been described as "the nontrivial extraction of implicit, previously unknown, and potentially useful information from data"[3] and "the science of extracting useful information from large data sets or databases".[4] It has close links with machine learning, statistics, databases and information theory. A traditional Data Mining process may involve all or some of the following steps: data collecting, data cleaning, data preparation, data integration, feature selection, feature engineering, model selection, model training and model evaluation. It may even require more than one iteration of some of the steps, since it is an exploratory process, susceptible to fail if not properly done. In the following sections we describe the relevant decisions we made in some of the most critical data mining steps.

II.A. Feature selection

In our approach, the selection of the relevant features to use as model inputs was done manually, based on human expertise. The reason lies in the fact that the physical meaning of the data involved is known.
We start by noting that the major external heat source influencing VEX thermal behavior is Sun radiation, followed by the Albedo effect from Venus. To capture this in the model we use orbital and attitude data of the spacecraft, such as distances and angles of these celestial bodies with respect to the 3 spacecraft orthogonal axes. Additionally we use the predictions of eclipses, since in that situation Venus Express does not get direct radiation from Sun.
Second, we also want to capture the effect of some internal heat sources. To achieve so, we include the predictions of power consumption for five payloads (Aspera, Virtis, Spicav, VMC, MAG) and one radio transmitter. PFS and VERA payloads were simply discarded because their profile of utilization is completely stable (during the period of study they were either always ON or OFF). Note that we don't include battery usage and other eventual internal heat sources, since there are no predictions available. Finally we include the previously mentioned higher-level nominal attribute: the pointing type, which can describe the overall operational status of the spacecraft. At this point we have 16 raw input attributes for which we have predictions available in the Venus Express Mission Planning System (MPS), enumerated in table 2.

Parameter	**Sampling**
Distance to Sun	6 min
Angle VEX X-axis with Sun	6 min
Angle VEX Y-axis with Sun	6 min
Angle VEX Z-axis with Sun	6 min
Eclipse Flag	2 min
Distance to Venus	6 min
Angle X-axis Venus	6 min
Angle Y-axis Venus	6 min
Angle Z-axis Venus	6 min
Aspera Power Consumption	16 sec
Virtis Power Consumption	16 sec
Spicav Power Consumption	16 sec
VMC Power Consumption	16 sec
MAG Power Consumption	16 sec
Transmitter Power Consumption	16 sec
Pointing Type	not fix

Table 2. 16 raw input attributes for the machine learning model

II.B. Feature engineering

Although our training data is a set of time series, when we build a machine learning model, such as a decision tree, the instances are considered as independent events as their order is not taken into account. This is particularly bad for this problem, since we know that changes in temperatures are cumulative and depend on the recent history of heat sources. In other words, even if we knew exactly, at a given time instant, the behavior of all internal and external heat sources, this would not be enough to make a prediction about the heaters power consumption. We need also to have information about the heat that was accumulated in the recent past due to the same heat sources.

To capture this cumulative effect of heat we decided to introduce some basic knowledge from thermodynamics.

Newton's law of cooling, states that *the rate of heat loss of a body is proportional to the difference in temperatures between the body and the environment*:[11]

$$\frac{dT(t)}{dt} = -r(T - T_{env})$$

Solving the differential equation, by standard methods of integration and substitution of boundary conditions, we have:

$$T(t) = T_{env} + (T(0) - T_{env})e^{-rt}$$

The temperature follows an exponential decay process. Note however that we are not interested on a rigorous analytical study of the thermal behavior of every spacecraft thermostat. In fact, heat transfer in space, occurs fundamentally due to radiation and conduction, since convection may be absent.[5] Our goal here is just to introduce a pre-processing step that can help the machine learning model to capture the system behavior more easily. The basic idea is: instead of just using the electrical currents and radiation intensities, what if we pre-process those time series to simulate heat accumulation? This could be applied to payload and transmitter currents and also to sun radiation time series.

The computational procedure to transform the raw time series is actually quite simple. Basically we use a sliding-window operator in which the power consumption of an instrument at time t_0 will be accumulated, with exponentially decayed weight, until t_0 + m (time from which we assume we can already discard the influence of the electric current in the local temperature). The parameters r and m were estimated based on telemetry data of Virtis payload and a thermistor in close vicinity. As an approximation we used the same "cooling" parameters for every time series.

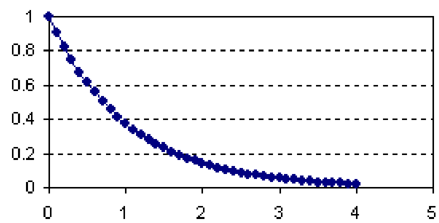

Figure 7: Exponential decay operator (time series following function e^{-rt})

With sun radiation, we additionally take into account the fact that radiation intensity decreases with the square of the distance VEX-Sun and it's completely absent during eclipses.

Input: Original signal S; Exponential decay operator Γ_k
Output: Exponential decayed time series T

foreach *time instant* t *in* S **do**
  foreach *time instant* d *in* Γ_k **do**
    $T[t+d] = T[t+d] + S[t] * \Gamma_k[t+d]$
  end
end
return T

Algorithm 1: Pseudo-code of algorithm to pre-process time series and simulate heat accumulation and exponential decay.

When we apply this preprocessing step Γ to a signal, we see that its shape changes considerably. More important, the time lag due to thermal inertia is captured. Note in figure 8 that the original signal (representing a payload power consumption) the time instant t_6 is exactly the same as time instant t_{10}, since the power consumption level is the same. But in fact, the thermal energy accumulated due to the electrical current is dramatically different, since at t_6 we are in a peak and at t_{10} we are at the lowest level.

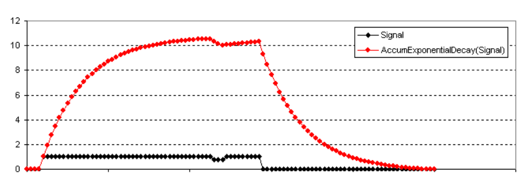

Figure 8: Example of application of the exponential decay operator to a simple time series.

Intuitively, thermostats respond to instantaneous temperature, not to instantaneous heat being produced. If we were not doing any pre-processing, mathematically we would have an *input-output mapping that was not a function*, since there could distinct output values for the same combination of inputs. This would mean that the statistical learning algorithms would discard important information, not only noise. The preprocessed input attributes are now listed in table 3.

Parameter	**Type**
Distance to Sun	numeric
Angle VEX X-axis with Sun	
Angle VEX Y-axis with Sun	
Angle VEX Z-axis with Sun	
Γ(Sun radiation)	
Distance to Venus	
Angle X-axis Venus	
Angle Y-axis Venus	
Angle Z-axis Venus	
Γ(Aspera Power)	
Γ(Virtis Power)	
Γ(Spicav Power)	
Γ(VMC Power)	
Γ(MAG Power)	
Γ(Transmitter Power)	
Pointing Type	nominal

Table 3. 16 pre-processed input attributes for the machine learning model. Γ() means exponential-decayed signals

II.C. Model Selection

The time series we want to predict is the sum of all LCL heater lines power consumptions. At this point we can follow two approaches: create a model to predict directly the total power consumption (see Figure 9), or create several models for each heater line and then combine the predictions to have the total (see Figure 10) .

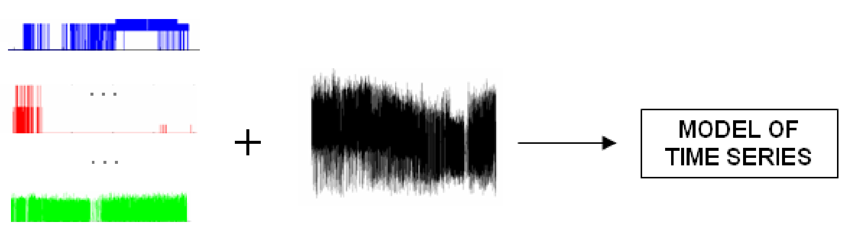

Figure 9: Possible approach: Sum all time series of heater lines power consumption and then create a model.

After trying the first approach we realized that we had at least one good reason to get better results by following the second alternative (bottom-up approach): in most LCLs the set of possible currents intensities is very small (telemetry indicates that the simpler ones have only 2 levels and the most complex around 50). In opposition, the time series for total power consumption can have thousands of possible levels (due to the combinatorial explosion generated by the sum) which makes it, in practice, a continuous variable.

This means that we could in principle split the big problem in smaller ones which are easier to solve. The question then is how to choose the appropriate model for each heater line and how to combine them. We deal with these questions in the next section.

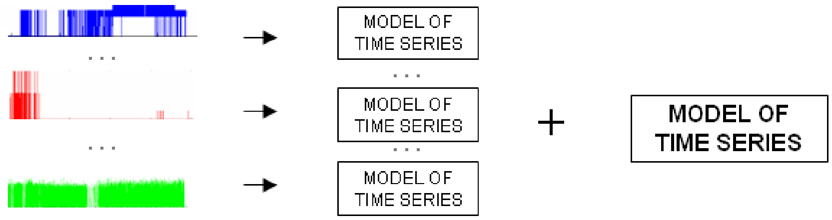

Figure 10: Possible approach: Create a model for each heater line and then combine them to predict total power consumption.

II.D. Machine Learning algorithms

Recall that each heater is controlled by a thermostat that sets an upper and lower temperature threshold. Additionally, we have seen that the set of possible states of the heater lines is finite and even small. This highly suggests that a decision tree can be a good learner to model the behavior of this system.

II.D.1. Decision Tree data structure

A decision tree is a predictive model which maps observations about an item to conclusions about the item's target value. Its interpretation is the following:

- each interior node corresponds to a variable;
- an arc to a child represents a possible value of that variable
- a leaf represents a possible value of target variable given the values of the variables represented by the path from the root.

In particular, we are using trees that support both nominal and real-valued attributes. There are several well studied algorithms to automatically learn decision trees from data, such as ID3 and C4.5.[8,9] Typically, the only information kept in the leaves after running the learning algorithm is its class or numerical value prediction. Here we have decided to follow a different approach. In order not to loose much information about the time series behavior, we introduced a change in the data structure and algorithm for decision trees: we keep histograms of the training data in the leaves of the tree. In this project we used the well known WEKA[12] data mining Java library.

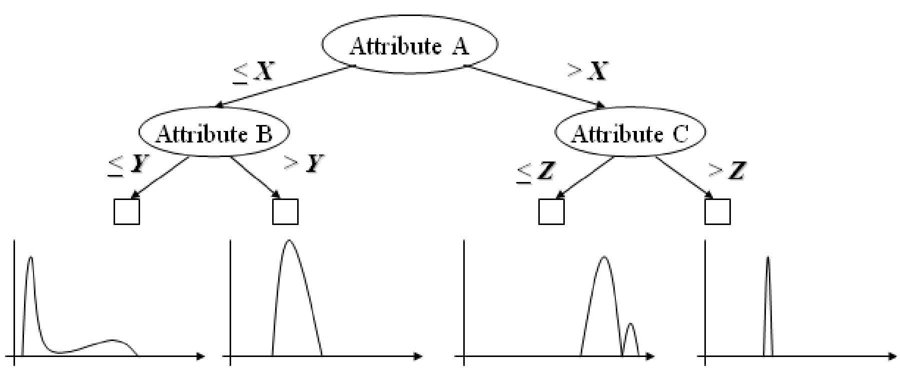

Figure 11: Decision Tree with data histograms of class attribute in the leaves.

These histograms should approximate the probability distribution function of the variable we want to predict.

Ideally, the histograms in the leaves would be Dirac-impulses, which would mean that we could predict a single value with high confidence. However, for complex and noisy domains, the histograms will likely spread over a wide range of possible values, which means we have to accept the uncertainty in our prediction.

II.D.2. Avoiding overfitting

Overfitting is the problem of using a statistical model with too many parameters with respect to the complexity of the data. For example, if we are modeling a linear process with some additive noise, we should only estimate two parameters: slope and y-intercept. If instead we use a n-degree polynomial, say with n=10, we will get a smaller error in the training dataset but a much bigger error in the testing dataset. This happens because the model "overfitted" to the noise in the data. In this sense, we can say that the model has learned the training data by heart and therefore generalized very poorly on new data.
When dealing with algorithms to generate decision trees, there are two basic approaches to avoid overfitting:[6]

- early-stopping - in which we stop growing the tree before it reaches the point where it perfectly classifies the training data,
- pruning - in which we allow the tree to overfit the data, and later we prune the tree.

In our case, we decided for a simple case of the first approach: limit the size of each decision tree (to height=3). This basically prevents the existence of leaves with very small number of data points falling on it, making the predictions more reliable.

II.D.3. Combining the sub-models

So far we have built a simple decision tree for each of the heater lines and the question now is how to combine them to have a prediction for the total power consumption. The trivial way would be to just sum up the prediction for the mean value of the power consumption for each heater. However, given that we are storing the data histograms we can perform a convolution operation to get the approximate probability distribution function for the sum of two random variables.[a] Recall that the convolution operation between two time discrete signals f and g is defined as follows:[7]

$$(f * g)(m) = \sum_n f(n)g(m-n)$$

Combining histograms by the convolution operation will make the resulting histogram wider, unless at least one of the histograms has width zero (Dirac impulse). For example, if the prediction for $LCLH_1$ was a uniform distribution in (10,12) Watt and $LCLH_2$ (7, 9) Watt, then the distribution for their sum would range from 17W to 21W. Note however that this distribution would no longer be uniform, since the value 19W is more likely to happen than 17W or 21W.

II.D.4. Making predictions

The fact that we keep data histograms in the leaves of the decision trees, allows us to predict not only the average value of the time series, but also upper or lower boundaries at each time instant.

We can for example extract the 95% percentile value in order to, most-likely, have an over-estimation of the heater power consumption at a given point in time.

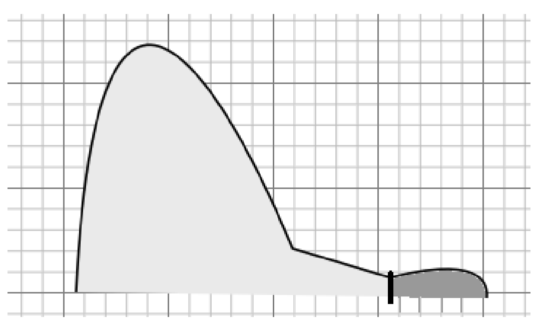

Figure 12: Histograms in the leaves allow the computation of upper boundaries.

[a]Note that by doing this we are assuming independence between the different models, which is not strictly true, but it turns out to be a fair approximation.

The motivation to exploit this capability of the model comes from the fact that it may be useful for operational purposes. In particular, the engineers using the Mission Planning System, may want to predict, with some level of confidence, the *worst-case* scenario for power consumption instead of the *most-likely* scenario. Note that the common practice to compute such a *worst-case* estimation is simply to add or multiply by a *security factor*. However, our approach is more informed because the added security margin is not constant, it depends on the histogram shape of each prediction.

Having said that, we now return to the scenario in which we compute the average values of the histograms, since they will give us the estimations that better minimize the prediction error. In the next section we discuss the quantitative evaluation of the model and its results.

II.E. Model Evaluation

To test our approach we split the dataset into two non-overlapping sections. The training dataset, used to create the models, covers the period from the 11th of April 2006 until the 31th of October 2006, while the testing dataset, used to evaluate the model, goes from the 1st of November 2006 until the 18th of December 2006.
The quantitative results of our model, measuring the average relative error of the prediction, are described in the table 4.

Training Dataset	Testing Dataset
16.81%	19.76%

Table 4. Mean relative error of prediction model

Note that the performance in the testing dataset gives us an indication on how good the model is making predictions for new data. To have a visual idea of the quality of the predictions we plot the real data together with the response of our model in figure 13.

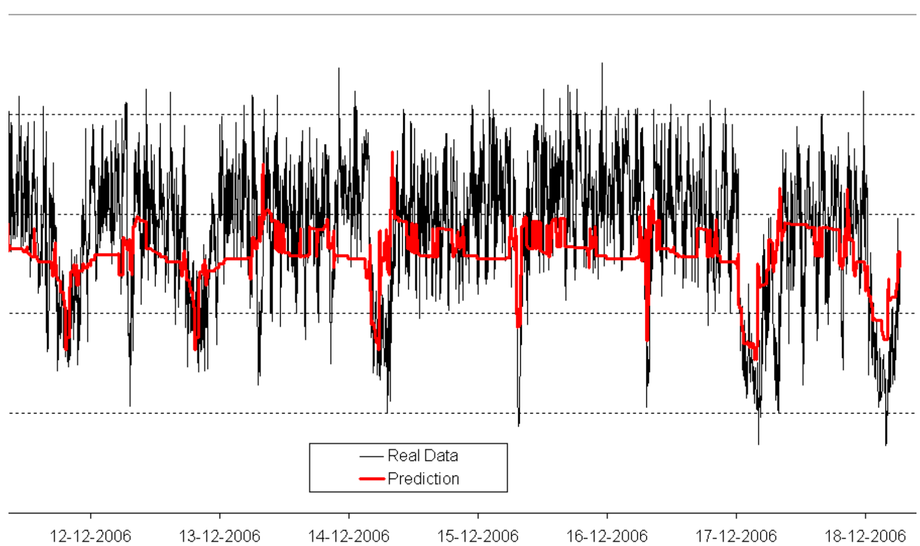

Figure 13: Plot of real data (black) and model prediction (red) in the last 6 days of the testing dataset.

Note that the high frequency changes of the signal are not learned, which is normal due to the complexity of the system we are modeling. Besides, for operational uses, it's enough for the VEX Flight Control Team to have a fair approximation of the average (or upper/lower bounds) and ignore the minute-level changes in the heater power consumption.
To have a rough idea about the weight of this high-frequency fluctuations, we did the simple exercise of smoothing the time series with a moving-average filter and then computed the average relative error with respect to the original signal. The values in table 5 indicate that although the results of our model are naturally worse than the real average of the data, the quality of our predictions is not far from what could be expected.

Training Dataset	Testing Dataset
13.31%	14.43%

Table 5. Mean Relative Error of Smoothed Data

III. Conclusions

This study started by proving the inadequacy of the simplistic *a-priori* model in use, and therefore the need for a more accurate alternative, based on time series sent by the spacecraft.

Often Operations teams deal with systems that are hard to model, but they are sitting on a load of flight telemetry data, which portrays the behaviour of the system under the most varied and real conditions, rather than those derived from engineering models or from ground tests. However, real life conditions are not always easy to isolate or correlate. Indeed, forecasting Venus Express thermal power consumption turned out to be a difficult task, due to the complex, dynamic, recurrent and partially non-observable domain.

Our Data Mining and Machine Learning approach proved to be a good way of dealing with this complexity and allowed us to get better approximations for the expected power consumption. Moreover, our model can help predicting worst-case scenarios, which can be of great use for operations. It has also been demonstrated that, when available, the introduction of some engineering expertise (like the one related to the thermal inertia behaviour) improves dramatically the quality of the empirical model.

The benefit of using this methodology to predict thermal behaviour is related to increasing flexibility in operations planning, since the current thermal constraints are defined in terms of a limited set of allowed envelopes representing standard science observation profiles (see[10]).

In a broader perspective, we showed that is possible to use new software technologies in the typically conservative space-domain, given that they can contribute to safety in operations and increase science return.

IV. Future Work

The next natural step in this project is the integration of the model within the Venus Express Mission Planning System. Another related task of interest would be the prediction of temperatures at specific points of the spacecraft. To achieve this, all the work done at the level of feature engineering could be re-used, although different modeling techniques may be required.

One can also envisage in the future that such methodologies are made available as standard library packages that can be plugged to a mission control (for data retrieval) and planning system (for resource usage estimation), allowing the Operations team to set up a model by configuring expected inputs/outputs, resolution algorithms, and additional engineering knowledge.

References

[1] Esa venus express official website. `http://www.esa.int/SPECIALS/Venus_Express/`.

[2] J Fabrega, T Schirmann, R Schmidt, and D McCoy. Venus express- first european mission to venus. *IAC-03 Q*, 2:06–54, 2003.

[3] W. J. Frawley, G. Piatetsky-Shapiro, and C. J. Matheus. Knowledge discovery in databases - an overview. *Ai Magazine*, 13:57–70, 1992.

[4] David Hand, Heikki Mannila, and Padhraic Smyth. *Principles of Data Mining*. The MIT Press, Cambridge, Massachusetts, 2001.

[5] Wiley J. Larson James R. Wertz. *Space Mission Analysis and Design*. Springer, 1999.

[6] Tom M. Mitchell. *Machine Learning*. Computer Science Series. McGraw-Hill International, 1997.

[7] J.G. Proakis, C.L. Nikias, C.M. Rader, F. Ling, M. Moonen, and I.K. Proudler. *Algorithms for Statistical Signal Processing*. Prentice Hall PTR Upper Saddle River, NJ, USA, 2001.

[8] JR Quinlan. Induction of decision trees. *Machine Learning*, 1(1):81–106, 1986.

[9] JR Quinlan. Learning decision tree classifiers. *ACM Computing Surveys (CSUR)*, 28(1):71–72, 1996.

[10] J. Schulster, B. Sousa, et al. Blowing hot and cold: Flight operations for thermal control on mars and venus express. In *Spaceops 2008*, number 123651. AIAA, 2008.

[11] Lindon C. Thomas. *Fundamentals of Heat Transfer*. Prentice-Hall, 1980.

[12] IH Witten and E Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2005.