Diplomarbeit

ILR-RFS DA 21-24

Comparative Data Analysis of Legacy Concurrent Engineering Studies

Hugh William Wessel

Zum
Erlangen des akademischen Grades

DIPLOMINGENIEUR
(Dipl.-Ing.)

Betreuer:	Dipl.-Ing. Christian Drobny Dipl.-Ing. Dominik Quantius
Verantwortlicher Hochschullehrer:	Prof. Dr. techn. Martin Tajmar
Tag der Einreichung:	29.01.2022
Erster Gutachter:	Prof. Dr. techn. Martin Tajmar
Zweiter Gutachter:	Dr.-Ing. Christian Bach

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die von mir dem Institut für Luft- und Raumfahrttechnik der Fakultät Maschinenwesen eingereichte Diplomarbeit zum Thema

Comparative Data Analysis of Legacy Concurrent Engineering Studies

selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Dresden, 29. Januar 2021

Hugh Wessel

Contents

Acknowledgements

For their assistance during the preparation of this thesis I would like to give special thanks to Phillip Fischer for helping me to gain a footing in the daunting task of parsing models of a complex and initially unknown data format, to Diana Peters for advice on finding insights within the torrents of generated data, to the specialists at the department of System Analysis Space Segment at the DLR Institute of Space Systems in Bremen for opening and facilitating access to their large resevoir of concurrent engineering data as well as their time and input on the brainstorming of the the project goals, and especially to Dominik Quantius for his regular support and unfailing readiness to answer my questions.

I would also like to thank Christian Drobny and Dr Christian Bach of the TU Dresden for their support in formulating a scope for the thesis and sorting out the administrative issues.

Abstract

To date, DLR has carried out more than 70 Concurrent Engineering (CE) studies in its Concurrent Engineering Facility (CEF). The vast archive of engineering model data generated during these studies offers the potential for obtaining a valuable insight into tendencies and correlations between engineering parameters which arise within the initial conceptual phase of space mission design. In many cases, the integration of Version Control Systems (VCS) has allowed access to a fine-grained log of the time-based evolution of a study model on a change-by-change basis.

However, due to the continuous variation in software packages and versions employed in the CEF over the years, it is currently difficult to make direct comparisons between the results of different CE studies. Additionally, as interpretation of the engineering model generally requires examining the accompanying natural-language documentation, a certain level of manual processing is unavoidable.

One area of interest for usage of such insights is the evaluation of Multidisciplinary Design Optimisation (MDO) methods, which requires some understanding of the progression of CE studies following a given CE process. Other potential derived benefits are improved methods for preemptive prediction of design feasibility and for generation of initial design inputs, as well as improvements to the design process itself.

The aim of this project is to develop a basis for semi-automated comparative analysis of legacy CE study data models by investigating a subset of DLR CE studies and associated engineering parameters and - where possible - to identify tendencies and correlations.

1 Introduction

1.1 Motivation

The digitalisation of the space sector and specifically the increased use of model-based approaches has increased the potential for the exchange and reuse of engineering knowledge and data, however the fast pace of adoption and variation of software solutions has lead to a fragmentation of data storage formats, making comparison between missions old and new a difficult and most often manually demanding task. In particular at the conceptual stage of space mission design it is of interest to form analogies to past missions, successful or unsuccessful, completed or in development, in order to avoid reinventing hard-won innovations or repeating the same mistakes. While work has been done so far towards processing existing static engineering documentation and data for the purposes of providing baselines for new studies, so far little work has been done towards observing the time based factor of how conceptual designs typically develop from the initial requirements into detailed specifications. Therefore there was a motivation to explore possibilities for observing the time-based evolution of engineering parameters during this initial stage as well as to consider ways of making comparisons between the conceptual process of former missions in order to determine the potential for using this information to augment or modify design processes, to assist in the implementation and benchmarking of design optimisation algorithms and to derive lessons-learned otherwise obscured when only considering the finished product.

1.2 Goals and Tasks

The goals of this thesis were the following:

- To deliver a review of previous research into the reuse of CE study model data
- To develop a conceptual basis for semi-automated comparative analysis of legacy CE studies and provide an initial proof-of-concept implementation
- To derive appropriate visualisation methods for interpreting CE data models
- To identify, where possible, trends and correlations in a subset of DLR CE data models, considering the time-based evolution of model parameters, including:

  - physical system parameters such as subsystem ratios for mass and power usage
  - design process parameters such as total number of iterations, number of iterations per session, number of components and margins

The following tasks were defined as a means to achieve these goals:

- Investigate the work done within the international community on the reuse of CE model data.
- Determine a list of engineering parameters and CE studies of interest for analysis.
- Develop a tool for extracting the desired data from the selected CE models.
- Define and evaluate potential visualisation methods for the interpretation of CE model data.
- Apply the chosen visualisation methods to the selected CE models and analyse the potential for identifying parameter tendencies and correlations.
- Assess the results of the comparative analysis and the efficacy of the applied visualisation methods
- Suggest possibilities for further work, where applicable with the usage of additional documentation such as study presentations and reports.

1.3 Document structure

Following this introduction, section 2 provides an overview of prerequisite topics for the remainder of the thesis. Firstly an overview of concurrent engineering is given, including an explanation of the role of software in the process and concurrent engineering as implemented at the DLR Concurrent Engineering Facility. Next, an introduction to multidisciplinary design optimisation is provided, presenting the investigative work carried out on the topic so far. Lastly, a summary is made of existing and emerging methods for the reuse of engineering model data.

Section 3 covers the process followed to achieve the goals defined in section 1.2 as well as the reasoning behind relevant scoping and implementation decisions. The description begins with the selection of concurrent engineering studies and parameters to be extracted for analysis. This is followed by an examination of the factors considered for the visualisation of extracted data, considering potential representations of the time-based aspect of model evolution, potential methods for measuring design convergence and possibilities for comparing the trends of analogous parameters in multiple studies. Finally, the data model and code structure developed to implement these process steps are explored, listing:

- the necessary assumptions with corresponding justifications
- a presentation of the metamodel employed for the mapping of study data to components and parameters of interest
- the code hierarchy for the characterisation and comparative analysis of concurrent engineering studies

Section 4 is concerned with a summary of the results of the data extraction process. The summary begins with a qualitative analysis of study parameters, detailing automatically extracted system modes and subsystem compositions on the example of the ASDR-II study. This is followed by an examination of model structure related parameters, including statistics on the complexity of models as well as the usage of capabilities to express modes and options among others.

In section 5, the results and outcomes of the thesis project are evaluated critically. Firstly the direct results of the individual qualitative analysis are discussed as well as the usage of the study data model. Then, the effectiveness of the developed solution for data visualisation and analysis is assessed, discussing the potential areas of application as well as the limits on the current implementation. Finally, an account of the difficulties encountered during the process is provided as well as remaining potential sources of error in the currently employed approach.

To conclude the thesis, section 6 summarises once again the content and outcomes of the thesis project, subsequently providing perspectives on the state of the problem space and potential avenues for further work on the topic of extraction and analysis of concurrent engineering data.

2 Background

2.1 Concurrent Engineering and the shared data model

2.1.1 The Concurrent Engineering process and exiting facilities

Concurrent Engineering is a collaborative methodology for producing and evaluating the feasibility of mission concepts in a significantly shorter time frame and with a higher degree of robustness than conventional mission design practices. A central characteristic of the concurrent engineering process is the parallel, i.e. concurrent, nature of interaction between the engineers. The Conconcurrent Engineering paradigm finds its place among the initial phases of a space project. The seven phases of a space project are defined by [1] as:

- Phase 0 - Mission analysis / needs identification
- Phase A - Feasibility
- Phase B - Preliminary Definition
- Phase C - Detailed Definition
- Phase D - Qualification and Production
- Phase E - Utilization
- Phase F - Disposal

The concurrent engineering paradigm has so far been employed in mission phases 0 through B however usage tends overwhelmingly to phase 0 and phase 0/A (beginning in phase 0 and continuing to phase A) studies [2].

A concurrent engineering team generally consists of a customer and coordinating team leader accompanied by domain specialists representing a selection of the following disciplines:

- Mission Analysis
- Science / Payload

- Communication and Data Handling (CDH)
- Attitude and Orbit Control System (AOCS)
- Propulsion
- On Board Computer (OBC)
- Software
- Structure / Configuration
- Mechanism
- Power
- Thermal
- Systems Engineering
- Product Assurance / Quality
- Assembly, Integration and Test (AIT)
- Project Management, Cost

The concurrent engineering process consists of four major stages in the following order:

1. Initiation Phase
2. Preparation Phase
3. Study Phase
4. Processing Phase

This thesis is concerned with the core "Study Phase". The study phase typically begins with kick-off presentations in which the mission drivers are presented by the study team leader, customer, mission analysis and systems engineer, from which an initial design can be iterated. Requirements are generally not fully formulated before the start of the study; the customer provides a set of high level business requirements which are refined throughout the study based on continuous interaction and communication of drivers and solutions with the engineers [3].

The domain experts are largely free to work independently during a concurrent engineering study, optimising their own subsystem following the appropriate driving equations and models, but are expected to communicate directly with other sub-system engineers on the numerous issues which have an impact on other domains. Indeed, as demonstrated by [4] and [3], the interdependencies between subsystems is both high-impact and complex. Coordination on subsystem interdependencies is therefore essential and is typically realised in two forms: direct exchange during domain rounds and ad-hoc consultation during non-moderated time. During domain rounds, each specialist presents the status of their respective area of responsibility and takes input from the other specialists. This leads to an initial number of changes to the design model during the session as well as introducing criteria which drive subsequent changes. During non-moderated time, it is up to subsystem engineers to initiate contact with others when deemed necessary and be prepared to receive feedback from others. The power of the concurrent engineering approach lies here within the flexibility and fluidity of interactions afforded by the close and simultaneous presence of all team members. Whereas conventional processes often involve a long serial chain of communication on single design points through slow and expressively restrictive channels such as email, the immediacy of contact during the concurrent engineering process opens up possibilities for more convenient parallelisation of design communication. The use of CE methodology as compared to traditional collaborative engineering methods has been associated with higher customer satisfaction, lower costs, increased quality and reduced need for redesign [5].

Concurrent engineering facilities have been opened in several different countries around the world, and during the literature research for this thesis, the research output of the following facilities was investigated:

- DLR Concurrent Engineering Facility (CEF) [5]
- NASA Jet Propulsion Laboratory - Team X [6]
- NASA Glenn Research Center - Compass Lab [7]
- EPFL Space Center Concurrent Design Facility (CDF) [8]

- Skoltech Concurrent Engineering Design Laboratory (CEDL) [9]
- TU Darmstadt Concurrent Engineering Lab (CEL) [10]

The results of this literature research are summarised in the subsequent sections where appropriate.

2.1.2 Software tools and data exchange

In order to realise the goals of concurrent development of a mission concept conforming to agreed requirements and constraints, it is crucial for the domain specialists to have a shared baseline for defining and modifying the design components and parameters. To this end, the centrepiece of any concurrent engineering infrastructure is the shared data model to which each specialist has controlled access according to their area of responsibility. While each domain specialist is almost certain to have their own local set of calculation models and tools for subsystem optimisation, it is only via the shared data model that the team can continuously communicate changes to their respective subset of the problem space which impact the design on the system level, short of verbally declaring these changes to each team member as is generally reserved for designated plenary sessions [11].

Much work has been undertaken towards defining formats for engineering data exchange. These have focused on developing intercompatible data exchange formats with the goal of easing the process of sharing model data between different organisations. A significant guideline for many data model implementations is provided by the ECSS technical memorandum "ECSS-E-TM-10-25A", which outlines model features recommended for the purposes of engineering design model data exchange [12].

On the user software side, there have been several MBSE software packages. The following software packages were found to be in active use and development:

- DLR - Virtual Satellite [13]
- Rhea Group - COMET (formerly CDP) [14]
- ESA - Open Concurrent Design Tool (OCDT) [15]
- Skoltech - CEDESK [16]

- Valispace [17]

Model data is generally stored in a central repository, with engineers editing the model via client applications which either require an explicit command or are capable of sending live updates automatically on any change. This parallel access however requires attention to permissions for editing to prohibit more than one user manipulating a value or structural element at the same time, which could either lead to confusion and errors, or at worst to corruption of the model data. The accepted compromise is to restrict each user to specific parts of the model, and in practice this does not represent a significant problem as components and subsystems generally map well to a single relevant discipline.

2.1.3 Concurrent Engineering at DLR

The DLR Concurrent Engineering Facility (CEF) in Bremen was founded in 2011 and has since provided a comprehensive facility for over 70 concurrent engineering studies. The in-house MBSE software package Virtual Satellite has been developed in tandem with the CEF and as of its fourth version is available as open source software [13]. DLR maintains a version of Virtual Satellite providing modelling features specifically for Concurrent Engineering Facilities, which builds on top of a core version that also serves as a base for the other offshoots, (Research, FDIR, TML) [13]. Virtual Satellite is built on top of the Eclipse Modelling Framework (EMF), a broad and powerful framework for developing both underlying model ontologies and applications in which these models can be created and manipulated [18]. Virtual Satellite 4 is delivered with a domain specific language, Concept, which allows for defining project-specific metamodels where the standard metamodel is insufficient. In addition, it supports the development of apps for in-study use. So far Virtual Satellite, has been used in over 25 concurrent engineering studies, of which 19 were carried out with versions 2 to 3 (both employing a similar base model) and 7 studies were held using the most recent Virtual Satellite 4 (employing a significantly different base model).

Figure 1 presents an example of a typical schedule for a concurrent engineering study in the DLR CEF. The typical scheduling events for a study carried out in the CEF can be divided into presentations (coloured green), off-line (coloured blue) and plenary (coloured red). Presentations include the project kick-off and final presentations, off-line scheduling events are non-moderated time, and plenary events refers to domain round sessions, short status reports and occasionally mid-study recaps. It should however be noted that these schedules are intended as preliminary agendas and are subject to change during the study, in which case the timetable is updated to reflect this for reference.

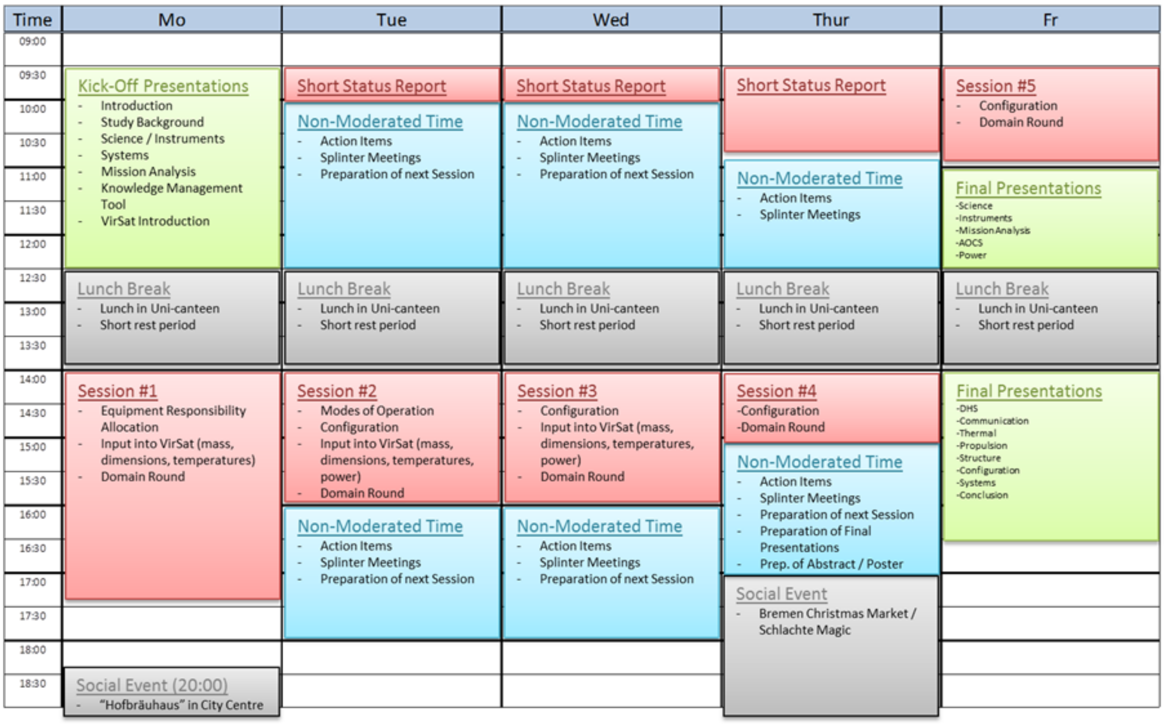


Figure 1: Timetable for the Aegis study

The studies carried out in the CEF have mainly been concerned with space missions, however feasibility studies have also been conducted focussing on other domains such as vertical farming, antartic stations and biohabitats [5]. Figure 2 shows an overview of the categories into which these studies can be divided.

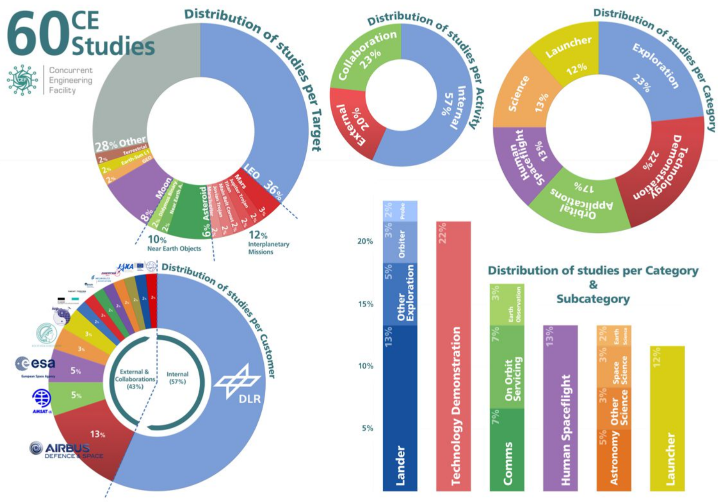


Figure 2: Statistics of past DLR studies from CEF Portfolio

2.2 Multidisciplinary Design Optimisation

2.2.1 General description and methods

Multidisciplinary Design Optimisation (MDO) is set of computational techniques employed with the goal of rapidly finding optimal system configurations. MDO frameworks exploit known relations between the disciplines involved in designing a system and attempts to optimise for the interests of each in parallel, while delivering a solution which is optimal on a system level. MDO has seen interest from several industries, such as aeronautical, railway and automotive, however the challenge of implementing MDO continues to present a barrier to practical applications in many cases [19].

2.2.2 Recommendations for implementing MDO at DLR CEF

A previous appraisal of MDO methods carried out within the CEF made recommendations for the introduction of MDO to DLR concurrent engineering activities, as well as defining necessary steps for realising their implementation [20]. From an analysis of several candidate methods and the requirements of the DLR CEF, it was suggested that Bi-level integrated system synthesis 2000 (BLISS-2000 [21]) was the most appropriate form of MDO for implementation.

Among the key prerequisites identified by the review is a better understanding of the dynamics of design variables on both the system and subsystem level, however to date little explicit data exists which can be referenced for this purpose. That is, in order to select an MDO architecture or method, it is necessary to be aware of coupling relations between subsystems and measure the extent to which all subsystems are coupled. It is hypothesised in this thesis that forming a basis for extracting time based data on the evolution of system and subsystem parameters from past CE studies and associating this with scheduled activities could provide input that can assist in the process of gaining this understanding. The reasoning is that optimisation tendencies observed during these studies could provide valuable insights on which future MDO implementations can be based.

2.3 State of knowledge reuse of engineering model data

Autonomous interpretation of engineering documentation Advancements in the fields of natural language processing among other artificial intelligence technologies has drawn attention to the potential of capitalising on the wealth of existing engineering data and documentation which has been generated within an increasingly digital context.

An ongoing project has focussed on developing an Design Engineering Assistant which is intended to leverage existing engineering documentation to derive insights and deliver these to engineers at the conceptual design phase [22]. Work on this project has already demonstrated potential to automatically interpret model data and provide structural insights in the form of knowledge graphs [23]. While this approach does not yet appear to have been applied to studying the time domain behaviour of engineering models, it could certainly be applicable to this task and make it possible to close the gaps in an otherwise automatable process of model comparison.

Application of statistical methods to conceptual design Work has been done at NASA Team X on a statistical model which can be used to provide feedback on the projected feasibility of design options before they are complete, using commonly stable relations such as mass fractions of subsystems which generally change little during the study after initial dimensioning and tend to end up as similar fraction of the final total mass [24]. This approach provides a means of recognising unfeasible designs early and avoiding wasting time developing these fully, however the approach does not take into account time-based trends of either the current or future design studies, which could potentially assist in making the projections more accurate.

Exchange of design model data Research which has been carried out at DLR has noted the importance on the intercompatability and transformation of model data formats in reusing designs from previous CE studies and missions [26] [27], however these topics still continue to pose a challenge, and it does not appear that the time-based aspect of versioned data models has yet been taken into consideration.

3 Procedure

3.1 Definition of scope

To begin the process of defining a scope for the analysis of DLR CE data, a brainstorming session was conducted with staff of the DLR Institute of Space Systems in Bremen and the DLR Institute of Data Science in Jena. The goals of this session were to:

- discuss previous work on similar topics at DLR
- identify drivers for selection of CE-studies for analysis
- discuss parameters of interest for comparative analysis
- formulate ideas for the goals for the subsequent analysis.

The outcome of this session had a significant influence on the goals and tasks as defined in section 1.2, however this scope continued to evolve during the initial stage of the thesis project as explained in the following sections.

3.1.1 Selection of CE studies

One main driver for selection of CE-Studies for comparative analysis in the brainstorming was determined to be the ability to categorise by target orbit or destination and mission type, however it was acknowledged that a significant factor would be limitations imposed by the diverse variety of model data formats employed during the lifetime of the CEF. As described in [2], the first 33 studies conducted within the CEF along with additional activities used a mixture of data models, including among others the Excel work books of the ESA Integrated Design Models (IDM-WB) and modified versions thereof, the “virtual System Editor ((v)Sys-ed)” from TU Munich, the Concurrent Design Platform (CDP), at the time developed by J-CDS, and the in-house Virtual Satellite (VirSat). Furthermore, while VirSat has become the standard for concurrent engineering in the CEF, the intent is to maintain the flexibility to continue using other data models as appropriate to the specific study, with the Open Concurrent Design Tool (OCDT) being a potential addition to the roster of software solutions.

While reliable data could not be obtained on the exact data models used in each study since those described in [2], it was noted that several studies used data models other than VirSat. However, as the majority of studies for which data could be obtained used VirSat and future studies are largely intended to use the same, it was decided that the analysis would include only studies using this platform. Nevertheless, the stated intention for a continuation in diversity of data model options in turn provided a driver to develop a baseline solution concept which, while initially concentrating on the analysed studies, could be extended to support other data formats in the future. Indeed, VirSat has seen a major revision since its first practical usage in the CEF, such that there are in fact two distinctly different data models - one corresponding to major versions 2 and 3, and another for version 4. Besides this, each sub-version has subtle differences in the baseline model applied to studies and version 4 even allows defining a specific metamodel structure for each study, however these factors were considered and could be assumed to be insignificant for the purposes of the analysis carried out in this thesis as elaborated later in (3.3.1). For this reason, it was decided to implement the proof-of-concept of the developed baseline solution such that it could bridge this gap in data model structure and provide insight into factors and limitations which would need to be considered in extending the solution to support further data models.

Thus, the scope of analysis was eventually reduced to the 22 CE studies listed in table 1, for which version-tracked model data could be obtained in either VirSat 2/3 or VirSat 4 format. For the studies which are confidential due to non-disclosure agreements (NDA) the number and name of the study has been shaded red.

For the purposes of the individual analysis demonstration, the ASDR-II study was selected as its use of the model allows a comprehensive application of data extraction and visualisation capabilities, and is the first VirSat study to be conducted with a team of experienced specialists. While the Aegis mission made extensive use of VirSat features, it is less representative of the usage of the data model in following studies. For the statistical analysis on model structure and procedure, all available studies were included, however in some results studies have been anonymised for reasons of confidentiality.

CEF Study No.	Study code name	Target	Type	Data model
26	Aegis	SSO; 700 km	Science / Astronomy	VirSat 2/3
31	ASDR-II	LEO (Debris)	Orbit Applications / On Orbit Servicing	VirSat 2/3
41	SHEFEXI-III	N/A	Technology Demonstration	VirSat 2/3
47	IoTA	N/A	Technology Demonstration	VirSat 2/3
49	OOS-RAV	LEO	Technology Demonstration	VirSat 2/3
50	Post-ISS Scenario I	LEO	Human Spaceflight	VirSat 2/3
51	TROJAN Lander	Trojan Asteroid	Exploration / Lander	VirSat 2/3
53	AIM/MASCOT 2	N/A	Exploration / Lander	VirSat 2/3
55	Post-ISS Scenario II	LEO	Human Spaceflight	VirSat 2/3
56	AIM-PALS	65803 Didymos	Exploration	VirSat 2/3
57	S2TEP	LEO	Technology Demonstration	VirSat 2/3
58	DEMOCRITOS	Mars / Europa	Technology Demonstration	VirSat 2/3
61	Post-ISS SNC	LEO	Human Spaceflight	VirSat 2/3
62	ROBEX RU	Moon	Exploration	VirSat 2/3
63	DARTS	Earth-Sun L2	Science	VirSat 2/3
65	QUANTUM	LEO	Technology Demonstration	VirSat 2/3
68	AMoCSIS	LEO	Technology Demonstration	VirSat 4
69	COFROS	LEO	Technology Demonstration	VirSat 4
70	FROST	LEO	Technology Demonstration	VirSat 4
71	N8	LEO	Science	VirSat 4
72	CO2Image	LEO	Science	VirSat 4
73	OSAS	LEO	Science	VirSat 4

Table 1: Table of available and version-tracked studies in VirSat 2/3 or 4 format

3.1.2 Choice of parameters for extraction

During the brainstorming session, input was provided by DLR CE specialists not only on parameters of interest for comparative analysis, but also on parameters which were expected to behave in certain ways during the study. The parameters of interest served as an initial driver for the extraction of data from study models, while the suggested expected behaviour of certain model aspects could be formulated into testable hypotheses.

As a starting point, the system mass budget was identified as an important and yet easily quantifiable parameter. [24] agrees with the claim that system mass has a strong relation to mission feasibility, firstly and directly through launcher mass requirements, and secondly and indirectly through the impact of system mass on mission costs. Modeling carried out by [25] suggests that, specifically for CubeSats, the mass has a statistically significant effect on costs. Further to this, the breakdown of the system mass over the subsystems of the spacecraft was indicated to be of interest. [24] highlights the importance of the mass of subsystems which are determined early in the design stage and statistically end up contributing a similar ratio of the total system mass in predicting the feasibility or infeasibility of a mission concept.

In addition to mass budget, the power usage and breakdown of subsystem contributions was also deemed to be of interest and is similarly a readily obtained parameter. While power usage itself does not directly affect the mission cost in the same manner as system mass, the power system itself can certainly be a significant driver of mission costs, and relations between the power system and the power usage of other subsystems can potentially impact other parameters as suggested by [24].

On the other hand, non-physical parameters relating to the structure of the design model were also identified as of interest for analysis. This included the number of components, number of iterations or model changes, the number of explored options, the number of subsystems and the number of modes. These parameters were deemed to be significant not only for the evaluation of CE studies, but also towards gaining a better understanding of how the features of the shared data model are used by engineering teams, both as a driver for requirements of the software platform and of the concurrent engineering process itself.

Besides these interests, as stated above, the expected behaviours of certain study aspects were expressed. Firstly, the payload mass was expected to be mostly given and to not change during the study, therefore having little impact on system design beyond its initial stated value. Secondly, the thermal system was stated to vary little during a given study and thus not be of much interest to an understanding of the evolution of the rest of the system. Finally, it was predicted that examination study models would provide little potential for comparing the level of detail and the usage of modes.

In summary, it was determined that initial investigation should concentrate on mass and power parameters and their distribution over subsystems, while structural parameters could be compared for additional insights into the CE process and associated data model usage, with certain parameters expected to behave in a certain way. How the evolution of these parameters was to be visualised and analysed is discussed in the following section 3.2.

3.2 Drivers for visualisation and analysis methods

3.2.1 Representation of time based design parameter evolution

The first factor to consider in visualising the time based evolution of engineering parameters was how to effectively visualise time itself within the context of a CE study. The absolute hour-by-hour time progression may initially seem a perfectly viable method for understanding the progression of the parameters over the scheduled study timeframe, however after little consideration the situation is revealed to be significantly more complex. Thus four different timescales were considered and weighed up against other in order to make an appropriate selection for application to the visualisations presented in section 4. Figures 3 to 6 present demonstrations of the considered timescales of absolute time, commit number, session block with intermediate values, and session block with final values respectively, which are elaborated in the subsequent paragraphs. For ease of comparison, the plots are grouped on the same page. The evolution of system mass during the Aegis study provides data for this example.

Firstly, it should be noted that it was decided to directly involve the scheduled activity blocks in the visualisation of parameter evolution. The reasoning behind this is that different types of activity as described in section 2.1.3 could be predicted to be associated with different behaviour of design parameters, while in almost all cases having potential to have an impact on overall trends. For instance, it would be expected that there would be little if no change during the final presentations, and most certainly not during initial presentations, so it would be considered neither surprising nor anomalous if no changes were to be observed over this period. On the other hand, it could be for example that long stretches of non-moderated time would be followed by a domain round in which numerous and significant model updates are made. In this case, it would be important to investigate whether this could have an impact on the optimisation performance of the team as a whole, as this could potentially indicate a delay of subsystem specialists in communicating parameter changes which could affect the optimisation potential of other subsystems. Considering these motivations, scheduling information was extracted from studies and integrated into time-based visualisations, as described below for the respective timescales. Where possible, colours for representing certain scheduling items were selected corresponding to those employed in the study documentation, however an exception was made in the case of overlap, where an additional colour was introduced. The chosen colour allocation is as follows:

- green for both kick-off and final presentations
- red for plenary sessions
- blue for non-moderated time
- yellow for status updates
- grey for lunch breaks

Out-of-hours time is not coloured, however a division is made at the boundary between days such that after-hours activity can be differentiated from early pre-schedule activity. Finally, a dividing line is included between all scheduling events for the case that two distinct items of the same type are depicted directly next to each other, as can occur to varying degrees in each of the considered timescales. For reference to the example presented in this section, see the time table for the aegis study shown in figure 1.

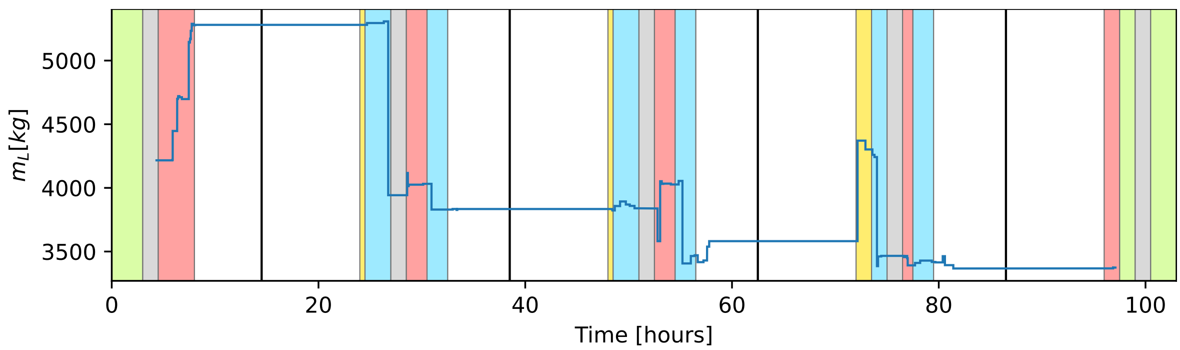


Figure 3: Plot demonstrating timescale by elapsed hours

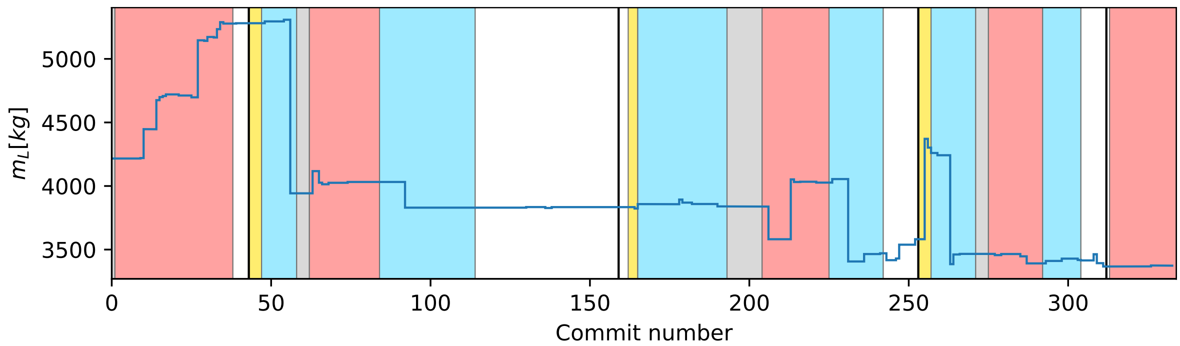


Figure 4: Plot demonstrating timescale by commit number

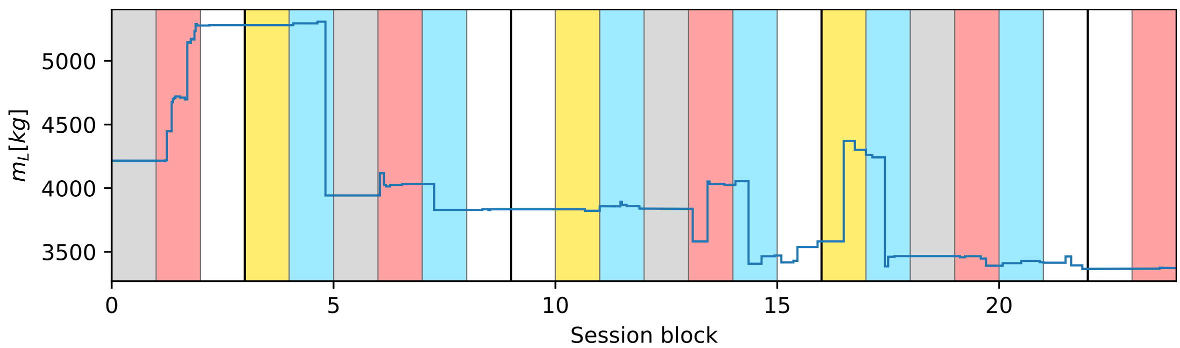


Figure 5: Plot demonstrating timescale by block with intermediate values

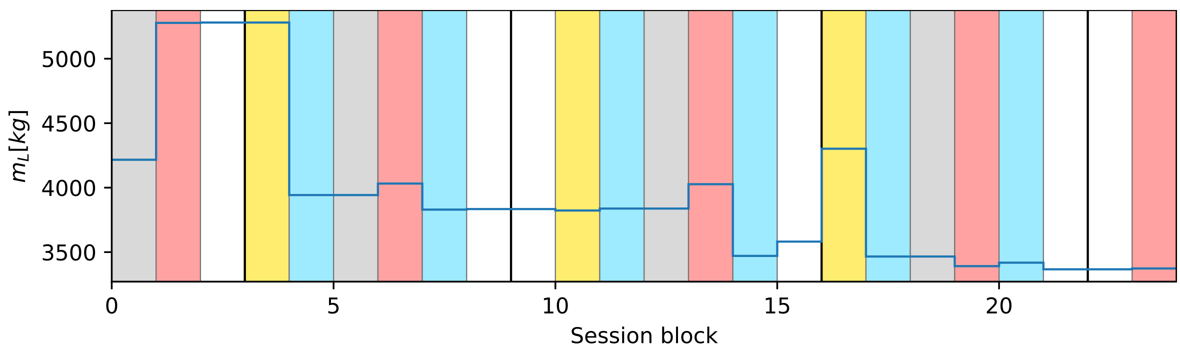


Figure 6: Plot demonstrating timescale by block with final values

As stated, the realtime number of elapsed hours was considered as a metric for study time progression. The immediate difficulty in using this definition however is the fact that no single study is carried out following the same agenda, or even following the agenda initially planned during study preparation. Due to the variability in types of scheduled blocks as defined in section 2.1.3, it would be insufficient to simply superimpose the value progressions of specific parameters over multiple studies, as behaviour could be strongly dependent on the type of activity taking place at a given moment. Furthermore, it later became evident through examination of the model update logs that the scheduled time as presented in documentation material was not even a reliable indicator for the actual hours of work, with evidence of model changes being made out-of-hours, both before and after official opening hours and during allocated lunch-breaks. Due to this it would not be representative to simply mask the data to scheduled time, excluding out-of-hours time, as this additional time could also be relevant to model progression, while not being readily quantifiable. It was concluded that while a real-time overview of parameter development could potentially provide insights into the study process, it would most likely not be of use in making comparisons between different CE studies. Nevertheless, figure 3 provides an example of this timescale. On first observation it appears that the majority of the diagram is empty space, however periods of apparent inactivity are inherently ambiguous, as model changes could have taken place which did not directly impact the plotted parameter.

The second considered timescale, using the commit number, shows more potential in representing periods of high activity, as now the width of a marked schedule item is proportional to the amount of commits, or changes, made to the model. For example, figure 4 reveals that the period in which the most changes were made was after-hours time on the second day of the study. On the other hand, this could lead to certain scheduling items being overrepresented, as the quantity of model changes does not necessarily correspond to the significance of the changes made, since larger changes in parameters may be made at once following longer discussions or simply following a longer period in which local changes were not committed to the shared data model.

Finally, the last two timescales represent each scheduling item in which changes were made as progressive steps of the study. The timescale demonstrated in figure 5 includes intermediate values spaced equally within each respective scheduling event, which gives an impression of how a parameter changed within a given block. The timescale demonstrated in figure 6 however takes a different approach, only presenting the final value of each respective scheduling block. This displays a view of the changes made in each scheduling item as single instantaneous changes. This may in fact be more representative of some CE activities, as often many changes are made almost in parallel due to group discussions such as domain rounds, so it makes sense to filter out sporadic intermediate variation and view these changes as single items. On the other hand, this view can potentially hide inter-subsystem interactions as it is no longer evident which subsystem was updated first, and which subsystems subsequently followed with changes.

Taking into consideration the advantages and disadvantages of each timescale visualisation, it was decided to make a pre-selection for later usage in the results presented in section 4. Both the timescale using commit number and that using schedule block number with final values were carried over as they both capture unique aspects of the parameter change behaviour - with the commit scale providing a measure of number of changes, and the block with final values scale providing a summarising overview of the outcomes of each schedule item. While this results in a loss of information about the absolute duration of scheduling items and the presence of items where no changes were made, the resulting plots can still be viewed alongside the study timetable to provide this missing context.

3.2.2 Qualitative analysis of parameter trends and correlations

Considering the scope of the goals set for this thesis as defined in section 1.2 as well as the scope of the studies available for analysis declared in section 3.1.1, it was evident that the dearth of data points of comparison would render the derivation of statistically significant quantitative models of the behaviour of engineering parameters during CE studies impossible. Therefore, this thesis concentrated more on defining a baseline framework for analysis of CE studies, employing a qualitative style of analysis to explore potential for insights to be gleaned from parameter progressions. Nonetheless, the qualitative observations made on the data of individual studies are simultaneously intended to provide indications for the immediate potential usage of the conceived visualisations as a tracing tool for future CE studies.

The aspects to be investigated are directly linked with motivations expressed in section

1.1 and elaborated in section 2: Firstly, the generation of test data towards the goal of implementing MDO techniques in the CEF, and secondly, capabilities for projecting the feasibility of a considered design option or reevaluating the implementation of CE processes. As the aspects of interest to each area of motivation are often overlapping, they are considered together in the following paragraphs.

Firstly, the convergence of system and subsystem parameters is of much interest, as this can assist in assessing the design maturity at the end of the study, or predicting whether certain maximum or minimum values will be exceeded or not during the study if a given design option is developed further. As an initial qualitative assessment, the convergence was measured by observing whether the parameter curves appeared to flatten towards a specific value. Convergence could also be assessed based on a plot of magnitudes of changes during the study, however as changes on a commit basis tend to be sporadic and widely spaced apart, it was decided for this thesis to produce change plots only on a schedule block basis. Besides simply measuring the convergence, it could also be of interest to relate this to initial values as well as activity indicating uncertainty at the beginning of the study.

Further to the detection of convergence in parameters, it is also of interest to identify interdependencies between the behaviour of parameter values which can be observed on a subsystem level. This can similarly be represented by the parameter changes by schedule block, however this demands special consideration in order to recognise changes significant to subsystems which themselves represent a small relative contribution, but may have a significant effect on subsystems with larger contributions. For this reason a plot representing such phenomena illustrates the changes in subsystems relative to their respective final value, as this is assumed to be representative of the theoretical value which would be taken into further design stages, provided that this value has converged sufficiently.

Finally as another point of interest, the percentage breakdown of subsystem contributions is to be included in the analysis, as this allows a clear impression of the final importance of optimising the parameters of specific subsystems.

3.2.3 Metrics for usage of the shared data model

The final stage of analysis handled in this thesis is the investigation of metrics related to the progression of the data model structure and the usage of data model features. As discussed in section 3.1.2, non-physical model parameters were deemed to be of interest for analysis, including number of components, number of iterations, number of explored options, number of subsystems and the number of modes.

For the number of components and number of iterations, these could be visualised in similar fashion to physically based parameters as described in the previous section.

As the number of options, number of subsystems and number of modes are generally by definition low and do not change significantly, these aspects are summarised as final value statistics. Whereas number of options is not expected to exceed 2, the remaining two parameters can be represented by frequency plots in order to gain an impression of the usage of these model features.

3.3 Data model and code structure

3.3.1 Assumptions

In order to realise the aims and goals for data extraction within the scope of this thesis, some simplifying assumptions were made in the development of the proof-of-concept tool described in the following sections. Moreover, in some cases assumptions were applied as a necessity to cover for apparent gaps or errors in model data. For all presented studies, multiple data points were manually verified, where possible with automated tests, however these assumptions would need to be considered if continuing development of the concept to either provide more reliable results, or to extend the concept to other model factors and data model formats.

The following assumptions were made relating to study model data:

- Where apparent errors in model data became apparent, such as cyclical references in equations or attempts to divide by zero, the associated commits were deemed to be invalid intermediate versions. Functionality was implemented to detect these circumstances during data extraction, which would otherwise have led to crashes, and to provide notification of the corresponding invalid commits, extracting no value for the concerned commit and skipping over to the next. For the few cases which triggered this notification, the associated model data could be reviewed manually and certain commit numbers were subsequently labeled as invalid and skipped entirely. Despite this, the measure of commit number would still count these revisions in the update timeline, however due to the infrequent nature of this issue this was assumed to have a negligible effect on results.

- In some cases, the units employed for analysed parameters were assumed to follow a common standard due to the complications involved in optionally processing units contained within the model based on the validity of the unit information data. For example, any power parameter considered was assumed to be expressed in Watt. As stated above, the values for studies appearing within the results were verified by comparing several datapoints with study documentation, and it was assumed then that parameter evaluations in between the manually verified values were also correctly calculated.

- A further example of unit assumption was percentages, which appear to have been previously defined without units due to percentage values only being used in certain types of equation. As it was not readily possible to differentiate this case without major modifications, a simplifying heuristic was defined that assumed any value processed in an equation known to represent a percentage (such as in margin calculations) was handled based on its size: in this case, any value below one was assumed to be an absolute factor to be applied directly, whereas any value above one was assumed to be a percentage and first divided by one hundred. This assumption was assumed to be valid as standards defined in DLR reference material declared that maturity margins would only be defined as one of three levels (5%, 10% and 20%), and system margins would be 20% in compliance with ECSS. While deviations from this standard were observed, no cases under 5% were noted.

- As mentioned in 3.1.1, various changes were made to the underlying base meta-model in between major versions of Virtual Satellite. This included addition and removal of model features and aspects and occasionally the modification of a datatype (for example the modification of a single precision floating point value to a double precision value). In the case of data type variations, the associated difference in accuracy was deemed to be negligible as all relevant numerical values to be processed were stored as text representations, and the effect of the change in precision of these values was deemed to be insignificant for the purposes of this thesis. In the case of the few modifications to model features, it became clear on manual examination that these features expressed design aspects which were neither defined in the shared data models of the considered studies nor relevant to the parameters extracted. Indeed for the most part, the base aspects of the model did not change and thus studies using the same major VirSat version were processed in the same way.

- The algorithms put in place to determine the inclusion of scheduling items in plots using the schedule block timescales assumed that no scheduled events would take place over the border between days. This assumption was easy to verify as all scheduling items were entered manually. For the case of out-of-hours changes on the other hand, a cut was made at the date boundary even if changes made after midnight could be assumed to be a continuation of an unofficial working session block. This is however not an issue for the studies presented in this thesis.
- Mode assignments which were extracted were assumed to be made on a system level rather than on a subsystem or equipment level. For the studies considered in this thesis this assignment could be verified manually, however this would have to be taken into consideration if extending support to other data models for which modes are defined on different levels.
- Some simplifying assumptions were made in deriving subsystem contributions to total system parameters, however these could be verified by comparing the sums of breakdowns with the total values and consulting study documentation.
- The prototype implementation developed for this thesis makes a number of further assumptions on the reliability of external services, in particular that of accessed study repositories, however as this is only a proof-of-concept it is not intended for the tool to be used as-is, rather these issues could be resolved through reimplementation of the developed concept.

3.3.2 Metamodel for description of CE Studies

In order to be able to compare CE studies on analogous parameters (e.g. total launch mass), or to provide a single report on the system or systems within a study in an autonomous manner, it was necessary to derive a format for expressing references to the appropriate parameters for which to extract values. Whereas the parameter elements in the model data may have obvious names such as “massLaunchMass” or “mass_launch”, without the use of natural language processing methods it is impossible to automatically detect which parameter in the model data corresponds to the desired parameter of interest. Furthermore, the name of a parameter can be changed at any point of the study or even the associated element deleted and replaced with a new element representing the same quantity, thus a single name or identifier for a given parameter alone does not suffice. The same applies to extracted modes or systems. For this reason it is necessary to provide information to the tool from which the correct identifier for a given parameter, mode or system can be determined at any revision number. The following section provides an overview of the data stored in the study metamodel files which encompass this information. See appendix A for the metadata files created for the analysed studies (confidential studies omitted).

CE Study metamodel Comprises a name for the study, information on the study storage repository, the software package used, whether the study is confidential, from which revision number the study model is valid (for data extraction purposes), the list of schedule items (see below schedule model), and finally the list of design options/configurations (see below system options model)

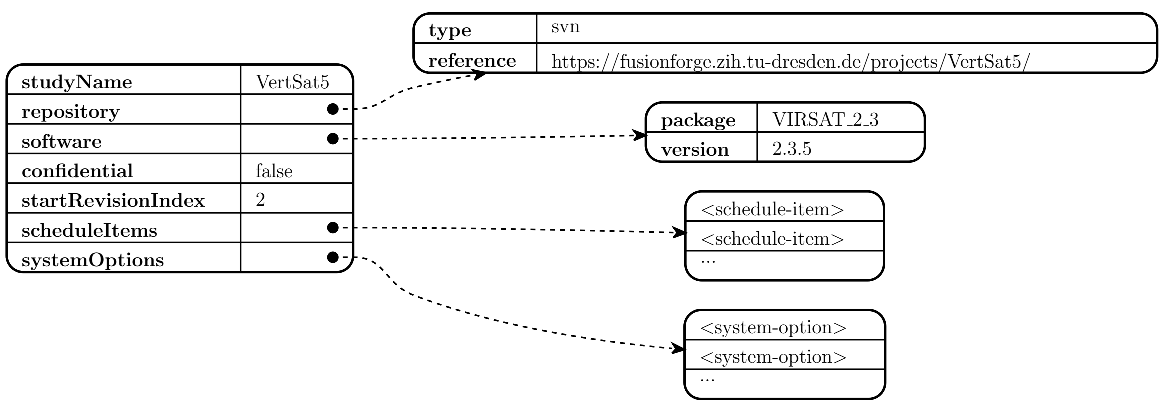


Figure 7: CE Study metamodel

Schedule metamodel The schedule items are defined as a list of events with a title for reference, a start and end time, and a type classification (presentation; plenary; offline; status update; lunch break)

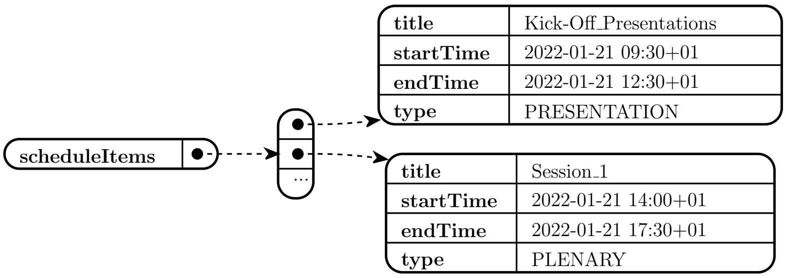


Figure 8: Schedule metamodel

System options metamodel The options considered within the study are each defined by a name, the modes applicable to the given option, the systems defined within this context, and a set of configuration parameters (not system specific) mapped to the parameter classification type, e.g. total delta V (see parameter of interest model below). [Note: this level of complexity is applicable to the studies OOS-RAV & DEMOCRITOS which both consider two options, plus debatably the post ISS missions I & II which at the end both have the "option" of breaking down by module rather than by subsystem]

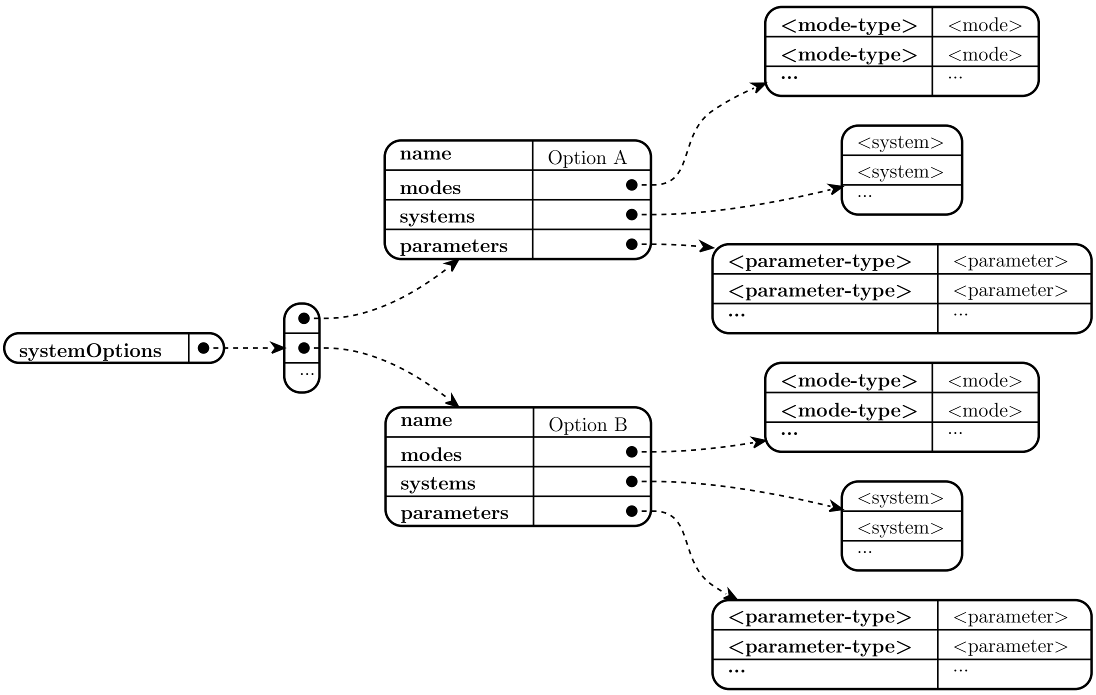


Figure 9: System options metamodel

Modes metamodel The modes of interest for comparative analysis are mapped to predefined types of mode (science mode; safety mode; etc) - currently only one per type supported. As mode elements can be created or replaced at any revision during the study while maintaining the same meaning, the identifiers for finding the specified modes (uuid) are indexed by the revision number from which they are applicable (e.g. mode for science is initially linked to an element with uuid of "93f...", but this element is removed and replaced at revision 10 with an element with uuid of "9a8..."). This enables the extraction of a consistent mode value history over the whole study, even if for some reason the identifier changes at any point.

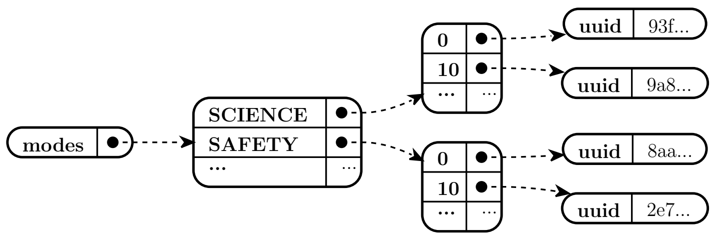


Figure 10: Modes metamodel

Systems metamodel The systems of interest for a given option are listed in an arbitrary order (although it is sensible to order higher level "systems of systems" first). Each system defines a name, a list of reference uuids analogous to the model for modes, whether the system is a system of systems, and a set of system parameters of interest mapped analogously to the configuration parameters.

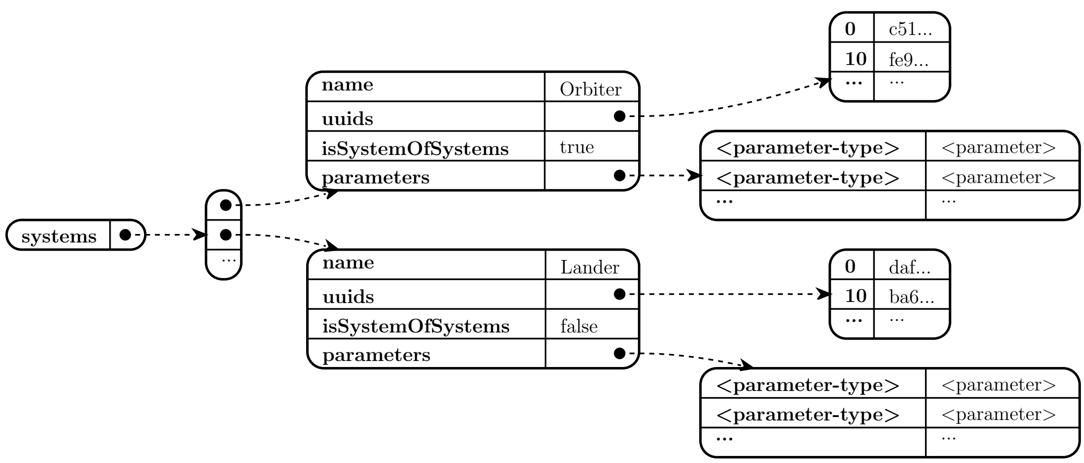


Figure 11: Systems metamodel

Parameters metamodel As stated above, each parameter of interest for comparative analysis is mapped to a parameter classification type in the same way as the modes (mass launch with all margins; average power with all margins; propellant mass) and also has a similar structure for dealing with potential element replacements during the study. Here however, each version of a parameter can optionally provide further references to other parameters representing constraint requirements (currently only maximum value). For revision ranges where the identifier is defined as null, the associated parameter is excluded from the data extraction process (otherwise the extraction process halts whenever the extraction tool expects to find a parameter at a given revision but fails to do so).

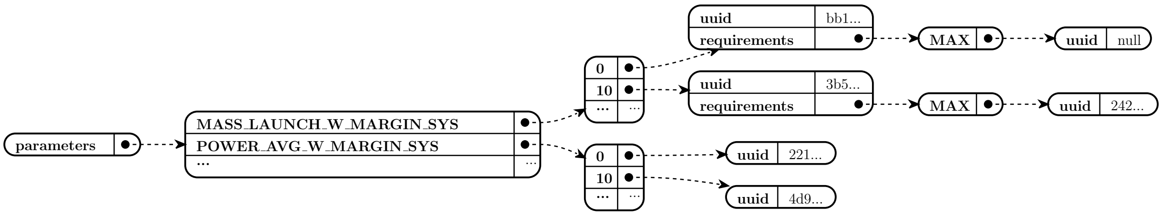


Figure 12: Systems metamodel

3.3.3 Data extraction tool hierarchy

In the following section, a high-level description is provided of the hierarchy of the Concurrent Engineering Study Data Analysis and Retrieval Tool (CES-DART). For a more technical insight into the tool, the source code can be found in Appendix B.

Figure 13 illustrates the hierarchy of the developed solution. The tool consists broadly of a pipeline to retrieve model data from a storage repository and store preprocessed data to local files, a model for accessing the data in concurrent engineering studies, a set of functions for analysing study data and an interface to access these functions.

The class 'CEStudy' is a format-independent interface for retrieving time-series data on study parameters, whereas the actual processing of format-specific model data files is delegated to an arbitrary "Study Model Handler", which can be exchanged with a model handler provided for any given format. Similarly the CEStudy class is independent of the type of storage repository containing the raw model data, instead allowing an arbitrary "Study Model Repository Handler" to be provided for retrieving the necessary files (so long as it is possible to reduce the study model data to a sequential list of revisions). When the value of a parameter at a given revision is requested via CEStudy, the model handler instructs the repository handler to fetch the corresponding revision from the study repository and store it in a local directory, from where

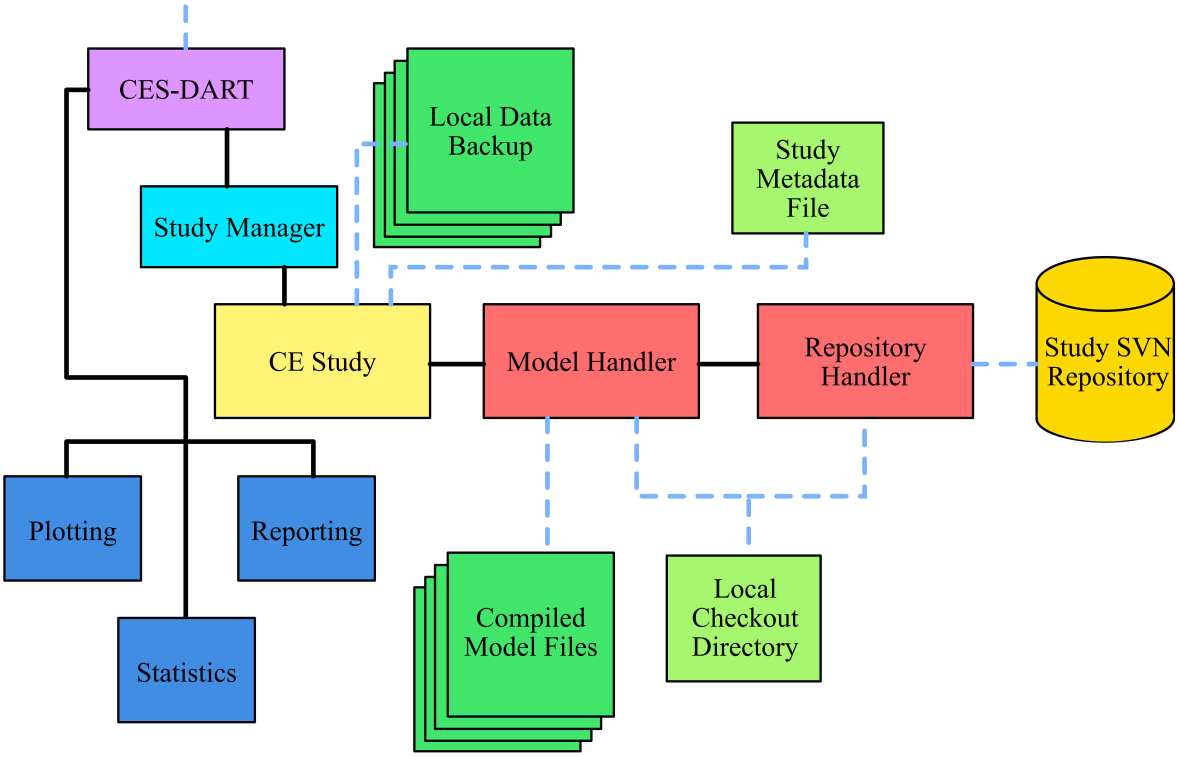


Figure 13: High-level diagram of CES-DART code hierarchy

the model handler pre-processes the relevant files and stores a compiled model file or directory for future requests of the same revision. The model handler subsequently parses the compiled model file or files to find and, if necessary, compute the requested parameter value and returns it to the CEStudy instance. As typically the history of a given parameter is requested more than once, the CEStudy will store a table of values in order to speed up successive requests for the same parameter. The information necessary to select the appropriate model and repository handlers and identify the desired parameter or system is expressed within the metadata file for the given study as defined in section 3.3.2.

For the purposes of generating parameter history plots, study statistics and study reports, a set of packages are provided containing the necessary functions. These can all be accessed through a single interface, whereby a 'Study manager' is responsible for retrieving a list of available studies and references to the corresponding metadata files.

4 Results

4.1 Individual analysis on example of the ASDR-II study

Figures 14 to 17 present the results of the study parameter extraction using the commit number as timescale, whereas figures 18 to 21 use the schedule block as timescale, showing only the values at the conclusion of each respective block. As noted in section 3, the regions of presentations, sessions, non-moderated time and status update blocks are coloured green, red, blue and yellow respectively. Those of scheduled lunch breaks are coloured grey, whereas out-of-hours activity is depicted with a white background, with a thick black line between days. Note also that due to the nature of the chosen timescales, only schedule blocks within which an update was made to the model appear. For ease of comparison, the figures have been grouped by timescale, with those using the same timescale being located on the same page. A selection of moments of interest are marked on each plot, with marking lines of the same appearance denoting the same moment in the respective timescale.

Figures 14 and 18 present the total launch mass with all margins, m_L, over the duration of the study, including the maximum allowable launch mass dicated by launcher requirements, while figures 16 and 20 present breakdowns, m_{sub}, of the same. Figures 15 and 19 present the default and mode values of the total average power usage with all margins, P_{avg}, over the duration of the study, while figures 17 and 21 present breakdowns, P_{sub}, of the default value.

For reference, the timetable for the ASDR-II study is also provided in figure 22.

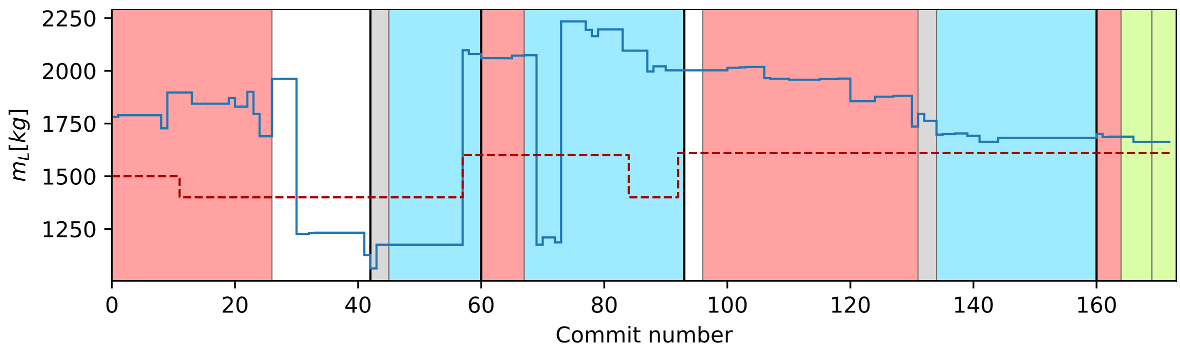


Figure 14: Plot of total launch mass with all margins (solid) and maximum allowable launch mass (dashed) by commit

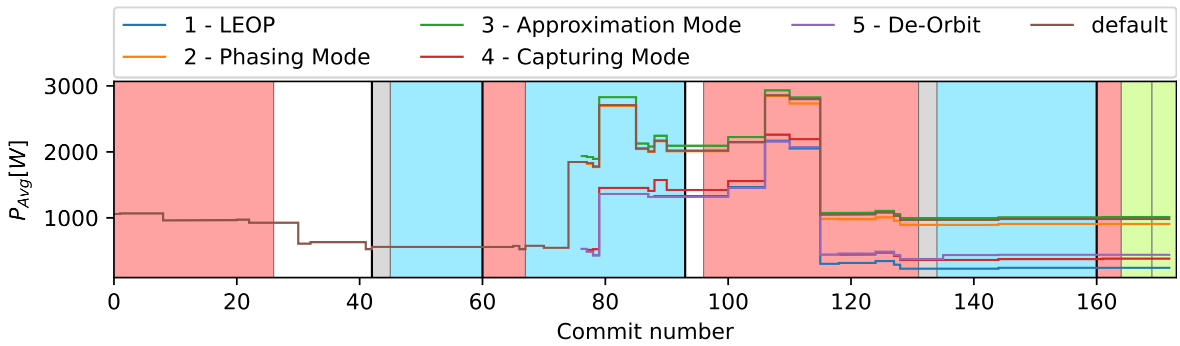


Figure 15: Plot of total average power usage mode values with all margins by commit

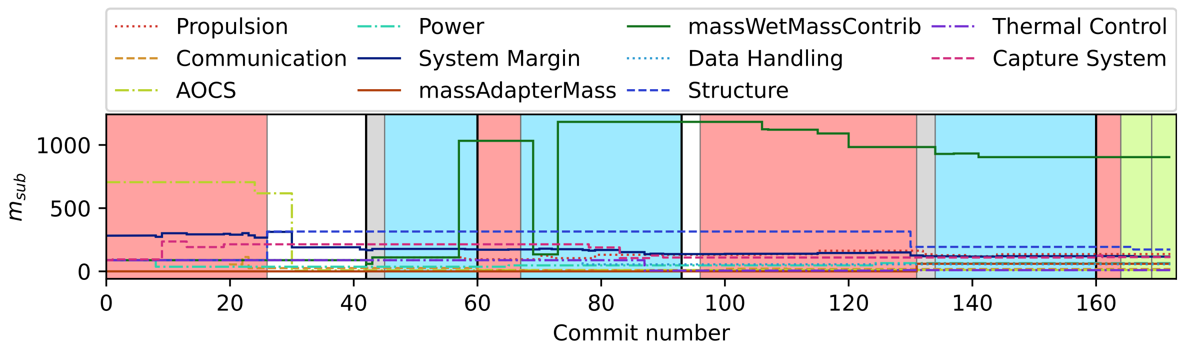


Figure 16: Plot of breakdown of launch mass by commit

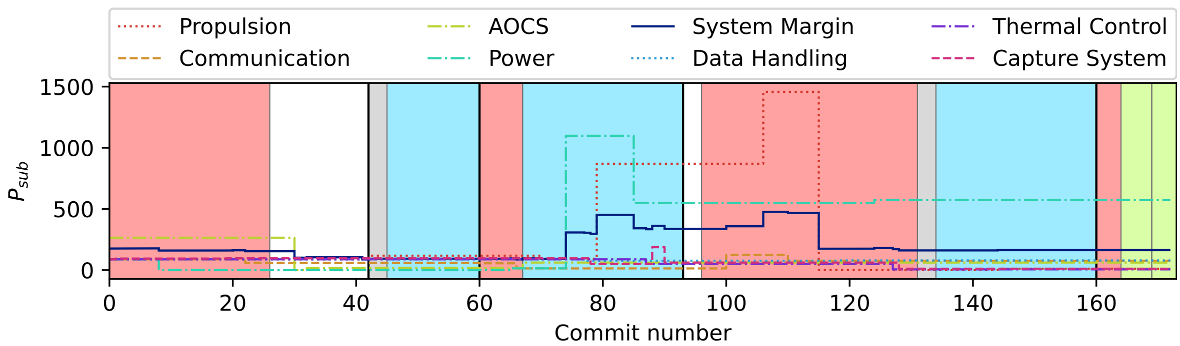


Figure 17: Plot of breakdown of average power usage by commit

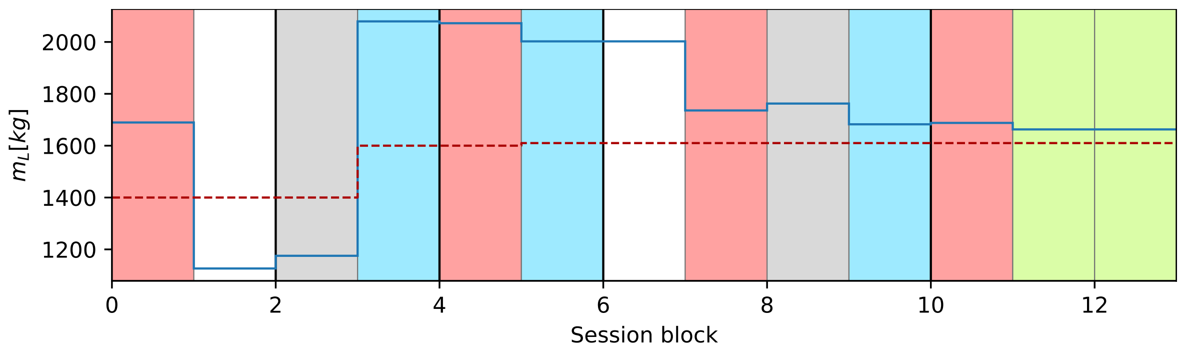


Figure 18: Plot of total launch mass with all margins (solid) and maximum allowable launch mass (dashed) by schedule block

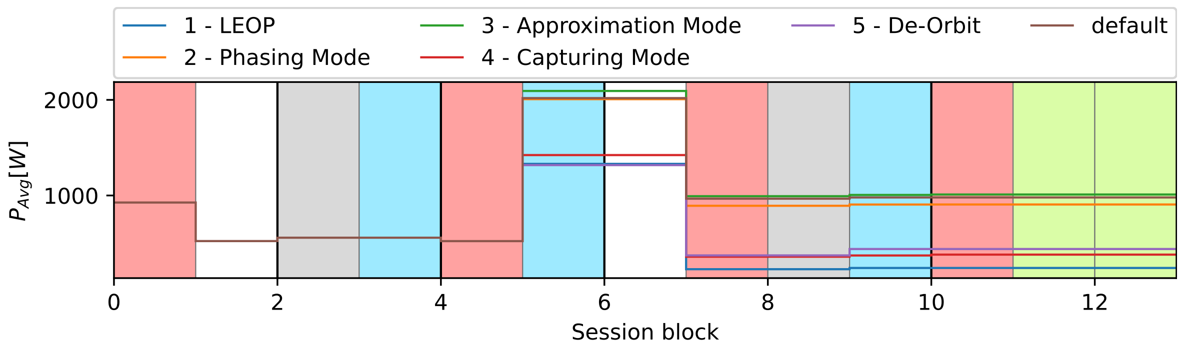


Figure 19: Plot of total average power usage mode values with all margins by schedule block

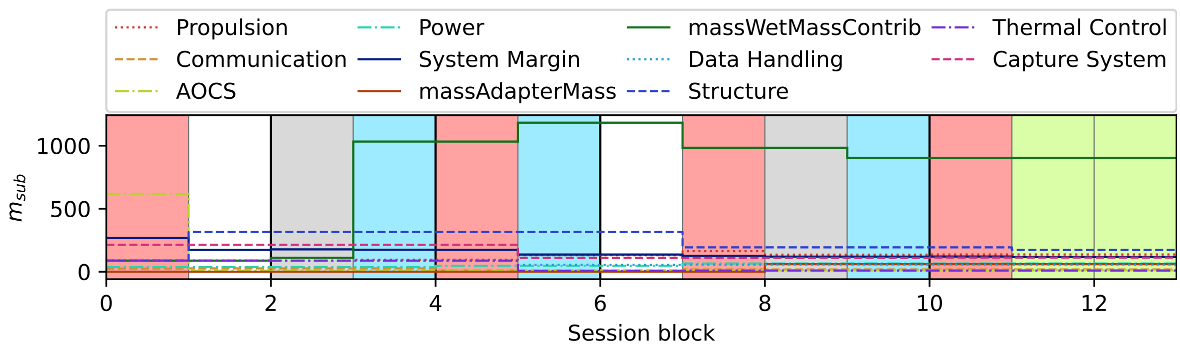


Figure 20: Plot of subsystem breakdown of launch mass by schedule block

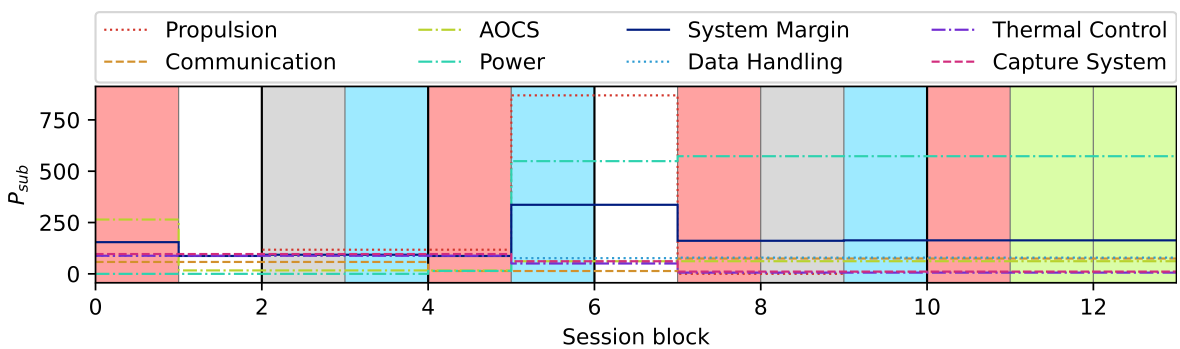
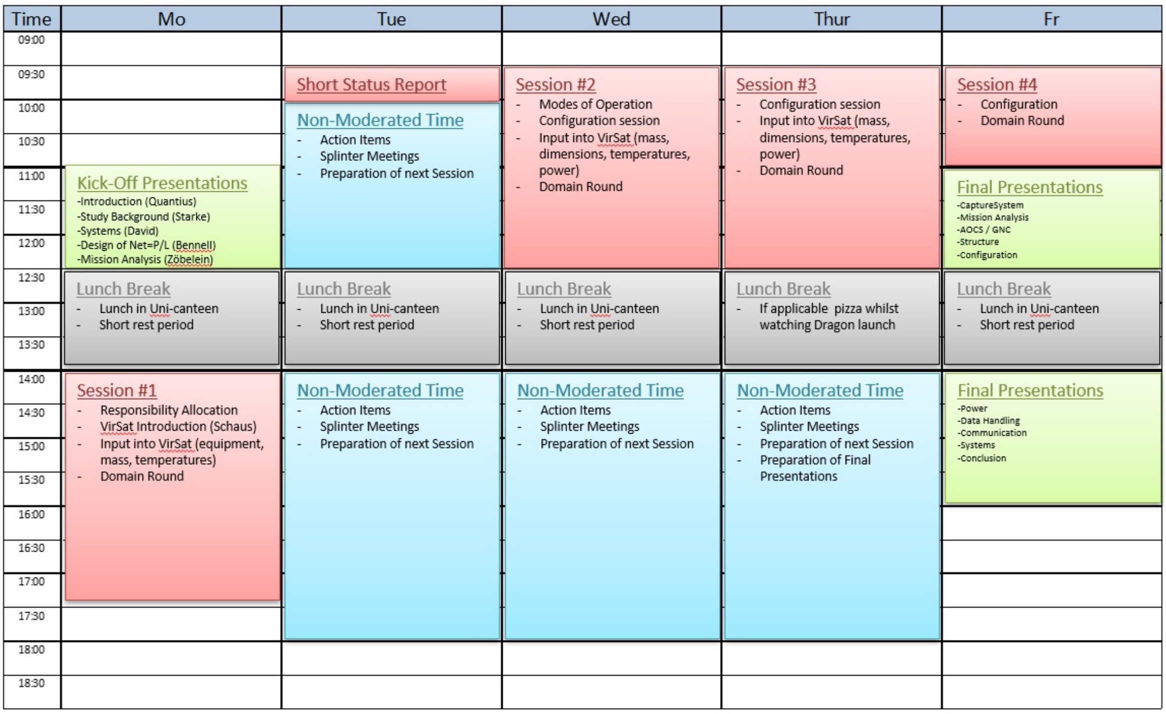


Figure 21: Plot of subsystem breakdown of average power usage by schedule block



Time	Mo	Tue	Wed	Thur	Fr
09:00					
09:30–10:00		Short Status Report	Session #2 - Modes of Operation - Configuration session - Input into VirSat (mass, dimensions, temperatures, power) - Domain Round	Session #3 - Configuration session - Input into VirSat (mass, dimensions, temperatures, power) - Domain Round	Session #4 - Configuration - Domain Round
10:00–11:00		Non-Moderated Time - Action Items - Splinter Meetings - Preparation of next Session			
11:00–12:30	Kick-Off Presentations -Introduction (Quantius) -Study Background (Starke) -Systems (David) -Design of Net=P/L (Bennell) -Mission Analysis (Zöbelein)				Final Presentations -CaptureSystem -Mission Analysis -AOCS / GNC -Structure -Configuration
12:30–14:00	Lunch Break - Lunch in Uni-canteen - Short rest period	Lunch Break - Lunch in Uni-canteen - Short rest period	Lunch Break - Lunch in Uni-canteen - Short rest period	Lunch Break - If applicable pizza whilst watching Dragon launch	Lunch Break - Lunch in Uni-canteen - Short rest period
14:00–18:00	Session #1 - Responsibility Allocation - VirSat Introduction (Schaus) - Input into VirSat (equipment, mass, temperatures) - Domain Round	Non-Moderated Time - Action Items - Splinter Meetings - Preparation of next Session	Non-Moderated Time - Action Items - Splinter Meetings - Preparation of next Session	Non-Moderated Time - Action Items - Splinter Meetings - Preparation of next Session - Preparation of Final Presentations	Final Presentations -Power -Data Handling -Communication -Systems -Conclusion
18:00					
18:30					

Figure 22: Timetable for the ASDR-II study

4.2 Structural and procedural parameters

4.2.1 Progression of model structure parameters

In figure 23, the number of components over study time in hours is presented for a set of studies. Figure 24 shows the number of commits over time for the same studies.

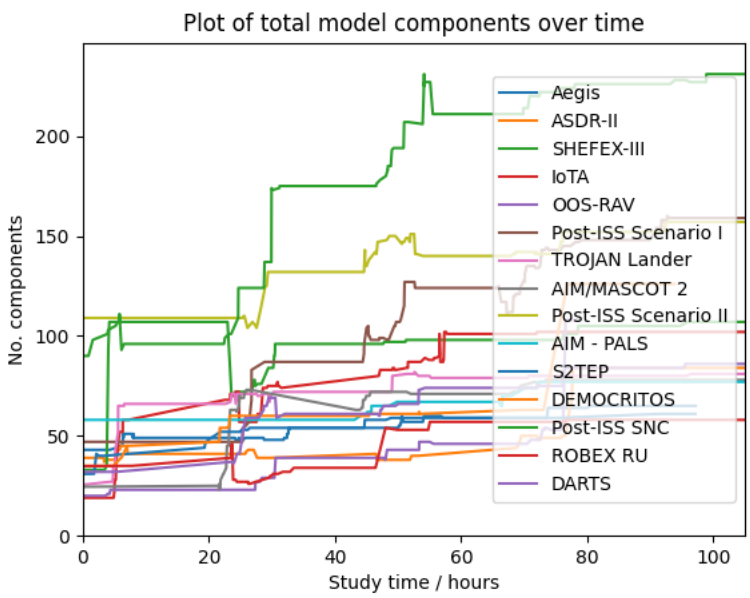


Figure 23: Plot of number of components over time for a set of studies

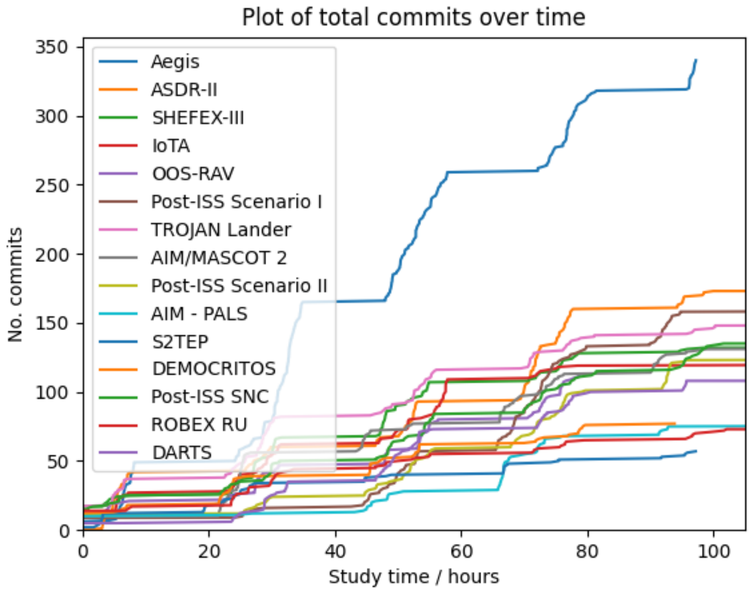


Figure 24: Plot of commit number over time for a set of studies

4.2.2 Model usage statistics

In this section, various statistics of model usage are presented.

- Root component count / options considered : 16 studies had only 1 option, whereas 6 studies had 2 options.
- Subsystem Count: Figure 25 presents a frequency chart of the number of subsystems.
- Modes count: FIgure 26 presents a frequency chart of the number of modes.

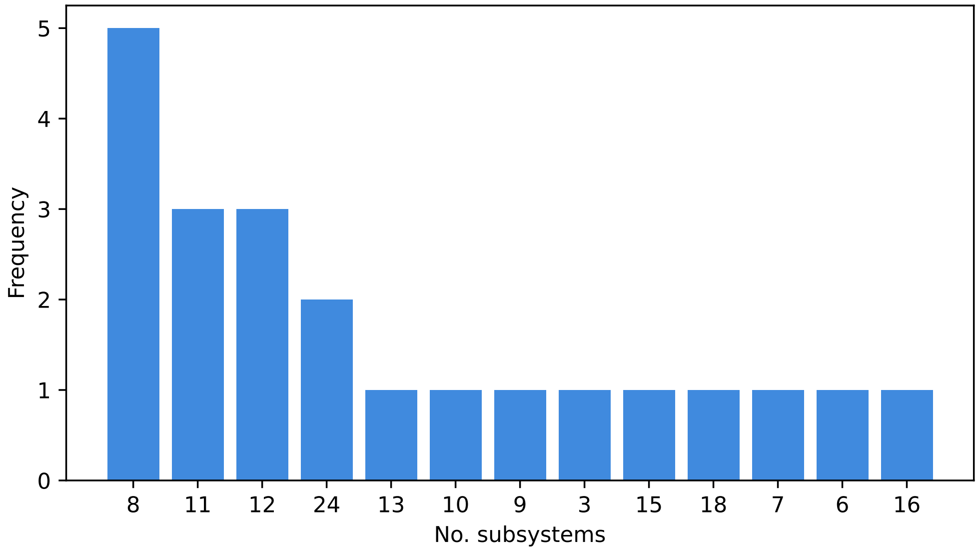


Figure 25: Frequency chart of number of subsystems in the model

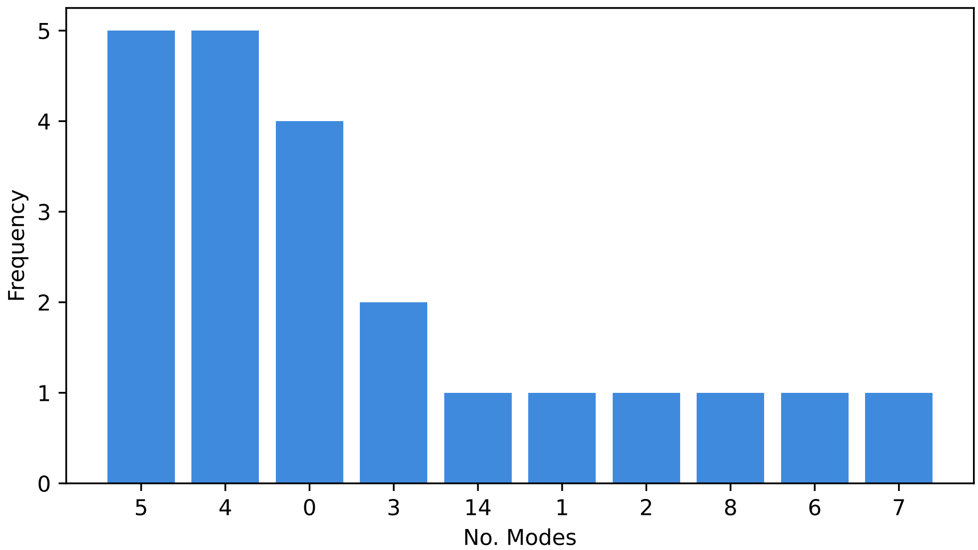


Figure 26: Frequency chart of the number of modes in the model

5 Discussion

5.1 Evaluation of results

5.1.1 Trends and correlations in study data

Starting with an examination of the results presented in section 4.1, it is immediately possible to observe a number of different procedural aspects. Firstly, examining figures 14 and 15, one can note differences between the behaviour of the total launch mass and the total average power usage. While the launch mass experiences many and significant changes from the initial commit during the first domain round through to the end of the third day, the power usage only begins to fluctuate significantly from the third day onwards. The fact that modal values for power usage only appear on the third day provide further indication that the power usage is initially given less attention, whereas there appears to have been much exploration on the mass. In the latter phase of the study, the launch mass appears to enter a more stable period of convergence, whereas the power use varies largely in the course of one non-moderated session and one domain round before arriving at a value similar to initial predictions and not changing significantly thereafter. Observing figures 18 and 19 this behaviour appears more explicit; while short-lived drops and rises are filtered out through the grouping of commits by schedule block, the same basic progression can be observed of an initial jump near the beginning followed by gradual convergence in the case of launch mass and a sudden fall in the case of average power usage. To a certain degree it appears that large drops in power usage may be associated with drops in the launch mass, however the inverse does not seem to apply for large increases.

Taking into consideration the maximum allowable launch mass, the convergence of the launch mass assumes more meaning. In this example, the system mass almost manages to go below the imposed launcher limit, however the failure to achieve this by the end of the study calls into question the feasibility of the explored option. It could nevertheless be that with one or two more days, a feasible conceptual design could have been found without drastic changes; potentially the time limit imposed by the final presentations may have led to less rigorous optimisation efforts during the last few schedule sessions in an attempt to wrap up the study and present the status of the design as-is.

Examining now the subsystem breakdowns of the launch mass and average power usage depicted in figures 16 and 17, it becomes apparent that most sudden large changes in a parameter can be attributed to one or a few subsystems. For the launch mass, the initially overestimated mass of the AOCS subsystem is corrected some time after scheduled hours before the end of the first day, with the sudden drastic rise on the second day being due to the introduction of the wet mass contribution into mass calculations (note that as breakdown contributions are automatically derived from the balancing equations defining the total value, an absence of a value for a contribution at a given moment indicates that this value was not included in calculations at that time). It could appear then that the wet mass contribution is the driving factor for optimisation, however a more logical explanation is that the wet mass is strongly driven by the sum of the other mass contributions, which can mostly be observed to decrease towards the close of the study. Importantly, this breakdown reveals that while the initially estimated total mass appears to have been a suitable prediction for the final value, this may have been more of a coincidence, given that the major mass contributions at either end are attributed to completely different entities. Taking a look at the power usage breakdown, AOCS is revealed to have been initially overestimated on power usage as well, whereas the sudden large increases in value on the third day being attributed to the propulsion and power subsystems. On the penultimate day, the power usage estimation of the propulsion subsystem drastically drops once again, however the power usage of the power subsystem in fact raises slightly and maintains the most significant contribution to total power usage through to study close-out. Once again, the block based timescale applied in figures 20 and 21 make these relations more explicit by filtering out sporadic and temporary changes during schedule items. Similarly, although the initial estimated average power use was very close to the final value at the end of the study, it is clear that the composition of subsystem contributions is radically different at both ends.

5.1.2 Usage of the shared data model

Observing figure 23, all studies unsurprisingly follow an upward trend, however there is variance in the rate of growth. Considering the initial component counts, it appears that this has a significant effect on the final value, whereas there are no studies which deviate significantly from their position between the other studies, except for temporary jumps.

As for the number of commits over time, there also appears to be some variance in the rate of commits between studies, however the patterns appear to be similar. While the timescale has the potential to confuse the impression of the tendencies due to periods between working days where no commits were made, it appears that in some cases, especially at the start the rise rate of commits is exponential, whereas this seems to flatten off towards the end. This would agree with the qualitative results in section 5.1.1, which indicate converging of parameters towards the end of the study. Indeed it is unlucky that teams would continue to make many and significant changes before the close of the study. One anomaly is represented by the Aegis study, for which commits were made more than twice as often than the next most committing study. This difference in tendency could be due to a demographical difference in the study participants; while other studies were carried out with trained engineers, the Aegis study consisted mainly of students. This seems to indicate that younger participants take a different approach to the concurrent engineering process, at least in their usage of the shared data model

Considering the results in section 4.2.2, it can be concluded that studies considering more than one option were a rarity. This could however hide options which were considered on paper but were not added to the study model, as this could cause unnecessary effort for a concept which is quickly discarded. Considering the number of subsystems, there is some variation, however the majority falls between 8 and 12 subsystems per system. It should be noted that in the cases where 24 subsystems were detected, this was due to the model containing multiple system options, explaining the unusally high number.

Finally considering the number of modes, it can be seen that they were applied in all but 4 of the 22 studies considered. The most common number of modes is between 5 and 4, whereas all other numbers are seen only once with the exception of 3.

In summary the statistics suggest a good use of the capability to represent subsystems and modes.

5.2 Assessment of outcomes

5.2.1 Efficacy of visualisation and analysis methods

The visualisations presented in section 4.1 succeed in providing a comprehensive impression of the evolution of the extracted parameters from several perspectives. Firstly, the commit based timescale is able to effectively communicate the amount and magnitude of individual changes, whereas the block based timescale is able to make the results of these iteration rounds explicit by removing noise caused by intermediate iterative fluctuations. From the initial example, it appears that tendencies and trends can indeed be identified in the behaviour of parameters and some expected relations between subsystems can be recognised. Where the representation is less effective is the relative inability to observe behaviour in subsystems which have provided similar and small contributions to the system total, however there is potential for using alternative heuristics to represent change such as plotting values relative to their final values or relative to a relevant maximum limit.

The visualisations in section 4.2 provide less clear cut results due to timescale issues, however it is still possible to observe general expected tendencies.

5.2.2 Potential for application

Equipped with visualisations of this type and the data behind them, there would appear to be potential for tracking and projecting parameters impacting design feasibility during future CE studies, as well as for gaining insights into CE procedures by examining past studies. The data extraction methods generated a large amount of data which was associated with relevant auxiliary data such as type and duration of activity and value limit requirements. This data could be of much use for future work in formulating MDO methods for CE studies, and the conceived metamodel provides an extendable baseline for extracting even more information from CE model data.

5.2.3 Concessions and limitations

While all aims could be realised to some degree, it must be acknowledged that there are limitations to the proposed solution as-is, and that there is still much work which can be done to solidify these goals.

Firstly, while the metamodel depicted in section 3.3.2 succeeds in providing a level of detail for the automatic extraction of parameter value histories in a way which complements ECSS specification, it stops short of mapping common subsystems to their counterparts in the study data model. That is to say that at this stage, the behaviours and interactions of subsystems cannot be comparatively analysed in the way now possible for system level parameters, although this would arguably provide more potential for characterising the dynamics of the design model over the course of a CE study.

Secondly, the available dataset currently supported by the suggested baseline solution for comparative analysis represents a very small sample size, from which it would not be possible to make statistically significant conclusions. The baseline solution itself may prove to have too limited a scope in its current state when confronted with further data model formats and would likely have to be modified to be more generally applicable.

Finally, the metamodel definition, data-extraction concept and qualitative analysis methods lack formal classification and thus could not be expected to immediately provide quantifiable results without first being more concretely specified. Moreover, the scope of the data-extraction tool was of such a breadth that there was subsequently not enough room for exploring all data which could potentially be derived from the considered studies.

5.3 Difficulties and errors

The difficulties encountered over the course of the thesis project stemmed primarily from the availability of CE data and ability to process the associated model files. Whereas the time-based extraction of parameter values initially seemed a trivial case of reading static values from a set of versioned files, the scope gradually considered more sources of error to be alleviated with a more comprehensive parsing of the model data. Principally, it was discovered at an intermediate stage that the static parameters values stored in text could not always be relied upon to provide up-to-date values for parameters dependent on other disciplines, due to the restrictions on changes to the model data by different specialists. For example, the system level parameters were mostly dependent on subsystem parameters, however if a subsystem engineer were to contribute a commit changing a driving value, this would not be reflected at a system level until the next update provided by the systems engineer, leading to sudden misleading leaps in the value history. Thus while on a schedule block basis this may not always be an issue providing the systems engineer updated the model last, the decision was made to implement parsing functions for evaluating the actual instantaneous values of parameters for correct representation of commit based visualisations. As the complexity of the data extraction increased, some unexpected tendencies in the model data were revealed which required deeper levels of specificity. As mentioned in section 3.3.2, the tendency for parameters or systems to occasionally be deleted at some stage during the study and be replaced with a new element representing the same entity, but carrying a different identifier introduced a necessity to handle different versions defined within the metadata files. Additionally issues such as equations containing references to the value they were intended to calculate as mentioned in 3.3.1 led to the need to manage more edge cases in the code. While this could be due to misinterpretation of the model data structure, examination of these situations did not reveal any alternative explanation.

Another unexpected difficulty was the identification of studies which could be used for analysis, as in most cases an explicit reference to the software or data model used let alone the specific version used could not be found. Moreover, it was not initially clear which models stored in the CEF versioned repository corresponded to which CE studies. While the original strategy employed required some level of trial and error, these circumstances motivated the migration to a custom model data parser and manual examination of the model data files. Eventually the solution applied was to identify candidate study models in the file storage repository and then review the available study documentation in detail in order to find matches.

Potential errors could be mistaken assumptions made in the parsing of study model data as well as erroneous intermediate processing of data, however through manual and automated verification processes this was deemed to be acceptable for the purposes of the implemented proof-of-concept. As more importance was placed on the general concept for comparative data extraction and qualitative analysis methods rather than deriving statistically significant tendencies, the importance of verifiably correct data was diminished. In the cases where static model features was considered however, these could be more easily determined with certainty due to the ease of manual verification.

6 Conclusions

6.1 Summary

In summary, the state-of-the-art in model reusage was investigated, and a data extraction and analysis tool was conceived that considers the data model from a different perspective, namely that of time. Visualisation methods were developed and assessed, and examples stemming from these efforts show that timescales based on number of model updates and schedule item were effective in communicating tendencies in engineering parameters. It was also observed that the drivers of trends in system parameters can be hidden within subsystem breakdowns, and that therefore reliable conclusions about progress in optimisation can not be made without this consideration. Finally it could be noted that features for expressing details such as system modes are well used, and that the rate of model updates displays convergent behaviour similar to system parameters.

6.2 Outlook and further work

Future work could concentrate on formalising the concept for comparative data analysis and extraction developed in this thesis. One immediate change could be the capability to associate subsystems in model data to commonly found subsystems in the way that modes and parameters have been characterised. This would allow for automatically comparing studies on specific subsystems such as AOCS, and for investigating correlations between them more closely.

Another concrete possibility includes extending visualisation capabilitites to represent the magnitudes of changes in system and subsystem parameters, as this could make correlations in parameter evolution more evident.

Finally, a significant area of interest would be extending the comparative analysis capabilities to CE studies beyond those carried out within the DLR CEF. While there are certainly many insights which can be gained from analysing internal studies, a broadening of the data set would be necessary in order to deliver statistically relevant results.

Bibliography

References

[1] "ECSS-M-ST-10C - Space project management: Project planning and implementation", European Cooperation for Space Standardization (2009)

[2] Braukhane, A., Quantius, D., Maiwald, V., Romberg, O.: "Statistics and Evaluation of 30+ Concurrent Engineering Studies at DLR. 5th International Workshop on System & Concurrent Engineering for Space Applications", SECESA, Lisbon, Portugal (2012)

[3] Gruschke-Lohse, A.: "Definition of a Guideline for the Development Process of a Spacecraft in Phase B based on Concurrent Engineering", Master Thesis, Bremen (2019)

[4] Ruhhammer, F.: "Untersuchung des Parameterraums einer Concurrent-Engineering-Studie", Bechelor Thesis, Aachen (2011)

[5] Martelo, A., Jahnke, S., Braukhane, A., Quantius, D., Maiwald, V., Romberg, O.: "Statistics and Evaluation of 60+ Concurrent Engineering Studies at DLR", DLR (2017)

[6] Case, K., Nash, A., Austin, A., Murphy, J.: "The Evolution of Team-X: 25 Years of Concurrent Engineering Design Experience", 2021 IEEE Aerospace Conference, IEEE, S. 1–6 (2021)

[7] Mcguire, M. L., Oleson, S. R., Sarver-Verhey, T. R.: "Concurrent Mission and Systems Design at NASA Glenn Research Center: The Origins of the Compass Team", NASA Glenn Research Center (2013)

[8] Ivanov, A., Borgeaud, M., Richard, M., Belloni, F., Trigueiro Baptista, A., Guevara, G., Enrique, M.: "Concurrent design facility at the Space Center EPFL", SECESA 2010 (2010)

[9] Golkar, A.: "Lessons learnt in the development of a Concurrent Engineering Infrastructure", INCOSE International Symposium, Vol. 26 No. 1, S. 1759–1769 (2016)

[10] Hoffmann, J., Bertrand, R., Ueltzhöffer, S.: "Concurrent engineering for ground segments and operations conceptual design", TU Darmstadt (2019)

[11] Fischer, P. M., Deshmukh, M., Maiwald, V., Quantius, D., Gomez, A. M., Gerndt, A.: "Conceptual data model: A foundation for successful concurrent engineering", Concurrent Engineering, Vol. 26 No. 1, S. 55–76 (2018)

[12] "ECSS-E-TM-10-25A - Engineering design model data exchange (CDF)", ECSS Secretariat (2010)

[13] DLR, Virtual Satellite, https://github.com/virtualsatellite/ (as of January 2022)

[14] RHEA Group, COMET-IME-Community-Edition, https://github.com/RHEAGROUP/COMET-IME-Community-Edition/ (as of January 2022)

[15] ESA, OCDT - Open Concurrent Design Tool, https://ocdt.esa.int/ (as of January 2022)

[16] Skoltech, CEDESK, https://cedesk.github.io/ (as of January 2022)

[17] Valispace, https://www.valispace.com/ (as of January 2022)

[18] Steinberg, D., Budinsky, F., Paternostro, M., Merks, E.: "EMF: Eclipse Modeling Framework", New Jersey, USA, Addison-Wesley (2011)

[19] Safavi, E., Tarkian, M., Ölvander, J., Nadali Najafabadi, H., Munjulury, R. C.: "Implementation of collaborative multidisciplinary design optimization for conceptual design of a complex engineering product", Concurrent Engineering, Vol. 24 No. 3, S. 251–265 (2016)

[20] Sanabria, O.: "Multi Disciplinary Design Optimization and appropriate recommendations to DLR Engineering Processes", Master Thesis. Bremen (2020)

[21] Sobieszczanski-Sobieski, J., Altus, T. D., Phillips, M., Sandusky, R.: "Bilevel Integrated System Synthesis for Concurrent and Distributed Processing", AIAA Journal, Vol. 41 No. 10, S. 1996–2003 (2003)

[22] Berquand, A., Murdaca, F., Riccardi, A., Soares, T., Genere, S., Brauer, N., Kumar, K.: "Artificial Intelligence for the Early Design Phases of Space Missions", 2019 IEEE Aerospace Conference, IEEE, S. 1–20 (2019)

[23] Audrey Berquand, Annalisa Riccardi: From Engineering Models to Knowledge Graph: Delivering New Insights Into Models, in: 9th International Systems & Concurrent Engineering for Space Applications Conference, SECESA 2020 (2020)

[24] Youmans, T. A., Hihn, J. M.: "Avoiding the impossible: Re-focusing a non-feasible mission 2-hrs into a 3-day engineering session", 2018 IEEE Aerospace Conference, IEEE, S. 1–14 (2018)

[25] Mrozinski, J., Saing, M., Hooke, M., Lumnah, A., Johnson, J.: "COMPACT KNN: Developing an Analogy-Based Cost Estimation Model for CubeSats", Jet Propulsion Laboratory, California Institute of Technology Pasadena, CA 91109, 2020 IEEE Aerospace Conference, IEEE, S. 1–9 (2020)

[26] Lange, C., Grundmann, J. T., Kretzenbacher, M., Fischer, P. M.: "Systematic reuse and platforming: Application examples for enhancing reuse with model-based systems engineering methods in space systems development", Concurrent Engineering, Vol. 26 No. 1, S. 77–92 (2018)

[27] Fischer, P. M., Schaus, V., Gerndt, A.: "Design Model Data Exchange Between Concurrent Engineering Facilities by Means of Model Transformation", Cypress, California, USA (2011)

Appendices

A Study Metadata

Please refer to the digital attachments provided. The study metadata files are contained within the directory "/CES-DART/Metadata/Studies/"

B Source code

Please refer to the digital attachments provided. The source code is contained within the directory "/CES-DART/Tools/". Additional scripts used to generate visual