Towards Using Functional Decomposition and Ensembles of Surrogate Models for Technology Selection in System Level Design

Citation for the original published paper (version of record):
Arjomandi Rad, M., Panarotto, M., Malmqvist, J. et al (2024). Towards Using Functional Decomposition and Ensembles of Surrogate Models for Technology Selection in System Level Design. DS 130: Proceedings of NordDesign 2024.
http://dx.doi.org/10.35199/NORDDESIGN2024.45

(article starts on next page)

Towards Using Functional Decomposition and Ensembles of Surrogate Models for Technology Selection in System Level Design

Mohammad Arjomandi Rad, Massimo Panarotto, Johan Malmqvist, Julian Martinsson Bonde, Kristina Wärmefjord, Ola Isaksson

Chalmers University of Technology, Gothenburg, Sweden

Abstract: Technology selection as part of the complex system design process involves multiple teams focusing on different subsystems, which in turn brings on interoperability challenges, extends design evaluations, and complicates design cycles. Utilizing function-mean modeling and the ensemble of surrogate model techniques, the paper reveals how low-level input parameters in the design can be instrumental in predicting higher-level performance outcomes. Applying the suggested framework to a case study from the space industry, surrogates trained on system level and component levels in a flow management system are shown to be generalizable. Exploring the methods for aggregating surrogates into ensembles shows how such an assembly can be built and utilized for navigating through the early fuzzy phases of design.

Keywords: Data-Driven Design, Conceptual Design, Functional Modeling, Product Architecture, Technology Selection, Surrogate Models

1. Introduction

Technology selection is the process of identifying, evaluating, and choosing technologies to realize a product's functional requirements and meet customer needs. However, in the early stages of design where technologies are typically selected by different teams, the vast size of the design space and interoperability issues can significantly affect the length and complexity of the design process. (Ulrich and Eppinger 2016). This complexity can obscure which technologies will most effectively meet project goals, necessitating a decision-making framework to guide the selection process effectively.

To avoid expensive redesigns and or having outdated concepts quickly, being able to analyze new technologies with data from previously designed concepts can be of great help. Designers constantly need to deal with continuously changing technologies that are used in a product architecture (Khadke and Gershenson 2009). Within a design process, analysis of existing historical data can be a valuable tool in evaluating new concepts because it can provide insights into successful design features. However, the integration of this data into the validation of new designs is still a gap that requires further research. Innovations often happen within a narrow scope, prioritizing detailed functionality, which can lead to integration problems when incorporated into a larger system (Iansiti 1995). Technologies that may underperform existing solutions can still offer significant advantages when viewed in the context of the entire system, highlighting the importance of a holistic evaluation approach (Panarotto, Isaksson, and Asp 2018). Current methods for generating and evaluating product concepts largely rely on the expertise of designers to define the design space during product development, a process that can be both lengthy and inefficient (Fazeli and Peng 2022). In essence, early-stage design is associated with reliance on limited-fidelity models and a lack of comprehensive performance data. This can lead to suboptimal design decisions and extended verification cycles and add unnecessary development lead time (Arjomandi Rad et al. 2022). The literature can benefit from new methods that allow high-fidelity models to be used already from early conceptual phases.

To address such obstacles, data analysis on product architecture that allows for the evaluation of concepts in the early stages of the design process can be an avenue to reduce complexity and accelerate design iterations. When designers are selecting technologies for different subsystems of a product, the possibility of predicting the performance of the system under study can significantly mitigate risks associated with technology integration. This study underscores the potential of integrating functional decomposition and an ensemble of surrogate models. We hypothesize that an early but rough approximation of the system's final performance can significantly widen the exploratory range of the design space while simultaneously reducing development costs. Therefore, the research question of this paper is: *To what extent can an ensemble of machine learning models that integrates system-level and component-level data offer enhanced support for technology selection in the early-stage design of complex engineered systems?* By enhancing the synergy between functional decomposition and the use of an ensemble of surrogate models, we aim to provide a more coherent framework for technology selection that leverages detailed component-level insight for system-level performance evaluation. The capability is showcased through a case product, Hall Effect Thruster (HET), which is a subsystem widely used to generate propulsion for satellites in space.

The remainder of the paper is structured as follows. After a brief literature review of similar and connected ideas in literature, the case study product is explained in the third section, where we also present product decomposition to present a better understanding of the different elements of the complex system. The results section presents performed data analysis and a ground idea for further analysis in ensemble models. The discussion section that follows next provides the implication of the findings and finally, a conclusion and some directions are suggested for future work section.

2. Background

2.1. Earlier work for Conceptual Design Space Exploration

In the early stages of product design, the evaluation of concepts predominantly employs low-fidelity models (often simple, abstract representations focusing on the core aspects of a concept). This is also known as the non-numerical approach in the literature, as opposed to numerical models in later stages which involve mathematical formulations and precise data for analysis and simulation) (Geng, Chu, and Zhang 2010) and its most known example is concept screening and scoring matrices by (Ulrich, Eppinger, and Yang 2008) that rank concepts based on a set of criteria and make selections in a two-stage methodology. The morphological matrix - originally developed by (Zwicky 1969) as a method for exploring all possible solutions to a multifunctional system - has been extended in modern applications to act as an optimization framework for aircraft conceptual design (Ölvander, Lundén, and Gavel 2009), which demonstrate its utility in generating and evaluation a broad spectrum of design alternatives. Other recent approaches to generate, assess, and prioritize design concepts more efficiently. For instance, the utilization of the Design Structure Matrix, Quality Function Deployment, and Extended Axiomatic Design has been exemplified in the creation of a hand rehabilitation device (Fazeli and Peng 2022). Another recent trend is synthesizing data from different development stages for use in the assessment of design concepts. For instance, the evaluation of design concepts that take into account broad user feedback and previous designs has been studied (J. Qi, Hu, and Peng 2021; Yuan, Marion, and Moghaddam 2021). Despite the existence of these supports, design space exploration, especially the task of technology selection, lacks consideration of how new technology could impact different aspects such as performance. The main reason is that low-fidelity models generally do not provide the detailed structure needed to effectively incorporate and take advantage of changes in product architecture driven by innovation.

Digital Twins have been used recently to address technology selection challenges in early conceptual phases (Tao et al. 2019) for market analysis, task clarification, and conceptual design. They have enabled real-time performance monitoring, predictive maintenance, and even the possibility of running "what-if" scenarios that simulate the effects of changes to individual modules on overall system performance. A modular-based flexible digital twin for factory design (Guo et al. 2019) and a scalable digital twin framework based on a novel adaptive ensemble surrogate mode are proposed (Lai et al. 2023). Adaptive digital twins architecture (Ogunsakin, Mehandjiev, and Marin 2023). A scalable digital platform for the use of digital twins in additive manufacturing (Scime, Singh, and Paquit 2022).

Excess data necessitates a range of common data-science activities to be carried out in analyses. Examples of these activities such as data acquisition methods, data harmonization methods, data storing, and integration to data processing methods (Liu et al. 2021; Q. Qi et al. 2021; Tao et al. 2022). Data science enables the analysis of vast amounts of data from various sources, including historical designs, sensor readings, and simulations, aiming to inform design decisions. By uncovering hidden relationships within data, we can predict performance, identify potential weaknesses, suggest design optimizations, and identify trends and patterns. In some cases, for tailoring designs to specific contexts, data from different operational environments have been fed into models to create designs optimized for those specific conditions.

2.3 Ensembles of Surrogate Models

Combining medium-quality models to yield better performance models (Barai and Reich 1999) is an early example of ensemble modeling approaches. Metamodels/surrogate models provide computationally efficient approximations of complex simulations and tests (Wang and Shan 2006). Aggregating multiple individual surrogate models to collectively provide improved predictions (Goel et al. 2007) is well accepted as an *Ensemble of surrogates*. This integration leverages the complementary strengths of different metamodeling algorithms (e.g., polynomial regression, support vector machines, and radial basis functions). Ensemble construction typically involves training metamodels on varied datasets or employing diverse modeling techniques to promote heterogeneity. From the technical point of integrating different models, several methods have been proposed such as *Bagging*: Metamodels are trained on random data subsets, with outputs often combined through averaging. *Boosting*: Sequential model construction focuses on improving predictions of where previous models performed poorly. *Stacking*: A higher-level meta-learner determines the optimal way to integrate predictions from the base metamodels (Mienye and Sun 2022).

The essence of ensemble methods is the way each method integrates different models. For example, some utilize global and local measures and an ensemble model in cases where only a small number of sample points are available (Chen et al. 2018). Others use a weighted average surrogate of response surface models on overall cross-validation error (Alizadeh et al. 2019). Using different models through a statistical system to improve the overall accuracy of the predictor can introduce more cost to the surrogate. It is proposed that (Lai et al. 2023) using ensembles of metamodels can improve prediction accuracy and reduce the additional cost in the ensemble through multicriteria model screening. Other than digital twins, ensembles have also been applied to analytical target cascading to tackle design optimization problems (Jiang et al. 2015) in literature. Most of the applications train multiple types of models on one problem and try out statistical aggregation of responses to achieve the goal of selecting the most adequate model (Stork et al. 2020). However, this paper proposes training one model per alternative technology rather than many models per one technology. This novel approach in utilizing the surrogates over product architecture and function-mean relation can push each model outside its training range but in total can capture system behavior.

3. Studied Case

Hall Effect Thrusters (HETs) are a class of electric propulsion devices that have become increasingly prevalent in satellite technology. HETs are inherently complex systems. They involve interactions between electromagnetic fields, plasma discharges, thermal management, and materials science. This complexity aligns well with the need for advanced tools and frameworks capable of handling a broad range of engineering systems. The main subsystem encompasses thruster Unit(s) where the acceleration of ionized propellant by electric and magnetic fields produces Thrust (mN). This ionization process is facilitated within a discharge chamber where a combination of a strong electric field and magnetic fields enables the ejection of high-velocity ions, which impart momentum to the spacecraft in the opposite direction (Boeuf 2017). Figure 1 encapsulates the operational workflow of a HET system in the form of a flow diagram (Kriebel 2002; Horizon 2019). This diagram helps to study different subsystems and their relations in terms of the flow of material, power, and signal.

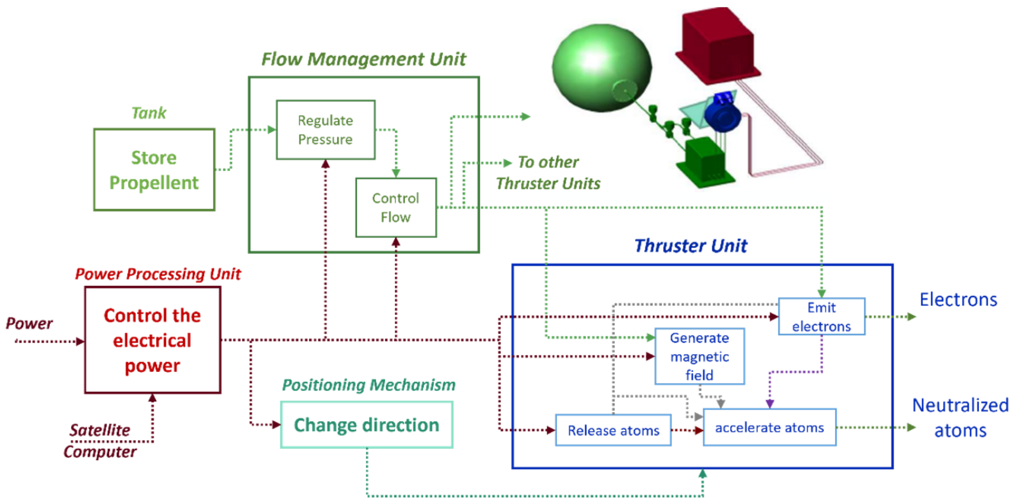

Figure 1. Flow diagram of a typical Hall Effect Thruster, The graphic on top right from (Horizon 2019)

For material flow, the figure shows a propellant tank that stores noble gases commonly Xenon, and recently Krypton. The flow management unit's primary responsibility is to regulate the flow of the propellant, maintaining a steady pressure and flow rate to ensure consistent thrust production (Kriebel 2002). It controls the balance between the propellant's storage conditions and the dynamic requirements of the thruster operation. This regulation is crucial for optimizing the thruster's efficiency and extending the operational life of the spacecraft (Di Cara and Del Amo 2019). As shown, HETs have well-defined modular subsystems which makes it ideal for experimenting with ensemble models combining both system-level and subsystem-level information.

A system engineering method useful to functionally decompose a system and study its subfunctions is the Enhanced Function Means tree (Schachinger and Johannesson 2000; Müller et al. 2019). Figure 2 shows a functionally decomposed HET tree that has "Keep/Change satellites in orbit" as its main function. This main function is then decomposed into several subfunctions such as "create thrust in space environment", "process power" and "ensure regulation of propellant". It should be noted that one can identify more subfunctions for such a broad main function. However, as shown in the lower levels of the tree, this study aims to examine two alternative means for one of the subfunctions "Ensure regulation of propellant", and therefore the subfunctions are kept uncluttered. The tree shows two alternative means for Flow Management Systems with differences in the level of integration of their sub-functions. In the first alternative pressure reduction and flow control are separate functions that are fulfilled by separate means. However, in the second alternative, these two functions are integrated and thus are fulfilled by one means. The illustrated alternative solution aims to fulfill the main function with fewer components, and control loops, which will result in having a simpler, lighter system which will essentially lead to less combined fault tolerance (Snyder et al. 2013).

The decomposed function means tree (where blue boxes are functions and yellows are means/solutions) in Figure 2 is given beside pictures of real subsystems of the provided means. The subsystem-level pictures belong to a thruster unit, a PPU unit, and a Flow Management System (FMS) (Vial 2023). The figure shows also two architectures for the two mentioned FMS alternatives that are composed of different numbers and compositions of valves, transducers, and heaters but deliver the same function (Kuiper 2023). Such alternatives can be used as technologies that can be replaced by each other because of having a shared interface, i.e. the same input/output relation. We investigate how much it is possible to predict high system-level performances such as Thrust using low-level valve data.

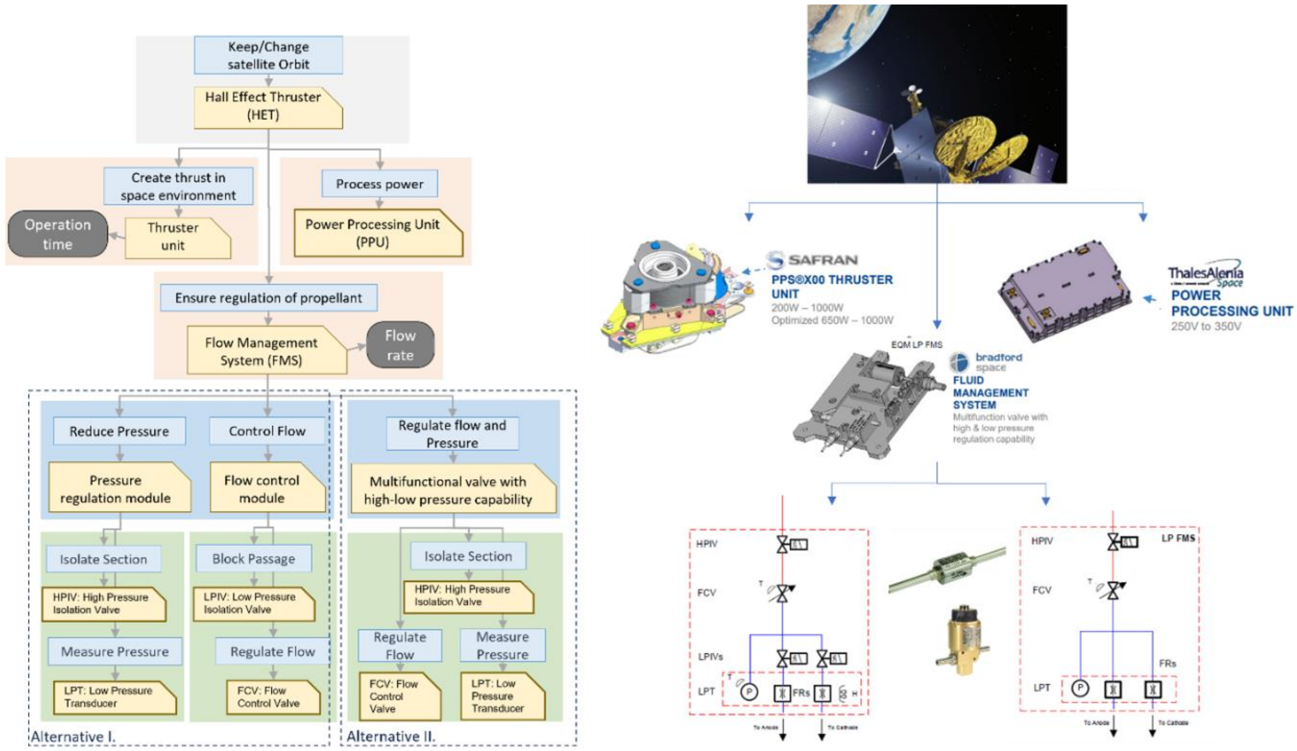

Figure 2. Functional decomposition of a Hall Effect Thruster with two alternatives for FMS, Right-side graphical elements derived from (Vial 2023; Kuiper 2023)

By mapping performance characteristics of diverse technological solutions onto a functional system architecture, designers gain a powerful tool for early-stage design exploration. This approach uncovers how specific technology choices directly impact system-level performance. Analyzing alternative means for fulfilling a single function within this framework enables rapid, data-driven comparison. Designers can seamlessly integrate and assess new technologies based on their predicted outcomes, streamlining the conceptual design phase.

To achieve the high-level function of "Keep/Change satellites in orbit" the system needs to create thrust and a large number of parameters can affect such performance metrics. Such parameters are imposed from different parts of the thruster unit. Figure 3 shows one node in the function tree that connects system-level performance (thrust) to low-level component performance (flow rate). In the top row of the figure, a HET is shown with thrust as a function of many parameters that are categorized as anode (listed in the figure), cathode, magnetic coil, and acceleration chamber has their own properties and parameters (Domínguez-Vázquez et al. 2022). One of the input parameters of the anode is the flow rate (mg/s) which is also considered a performance output in the lower hierarchy and can be determined with parameters such as inlet pressure and orifice diameter for the valve as shown at the bottom row of Figure 3. The figure also demonstrates two alternative solutions (previously shown in Figure 2) as *new* and *previous* manufacturing solutions. The higher level dataset of the system with flow rate as X axis and thrust as Y axis for the system and in the same way the subsystem level dataset for the valves at the bottom layer, the flow rate is the Y axis and is plotted in different colors for each orifice diameter. The idea is to use these datasets from different parts of the product architecture and train a series of machine learning models as an ensemble of surrogates in an integrated analysis.

Building an ensemble of metamodels over a functionally decomposed architecture has several advantages. One of the key benefits is that it allows us to use component-level data for decision-making at the system level and vice versa. This can help us scale our systems and components in terms of performance. Having a group of surrogates resembling the functional structure of the product can enable us to assess different technologies' performances in a plug-in way, given that we have the surrogate for the new technology. This modularity of the method enables different means that address the same function to be replaced through their shared interface which in this case will be their input/output. An ensemble will allow scaling the solution space through changes that can be driven by variations of input parameters of the system.

While the predictive models are limited due to the availability of comprehensive operational data on physical HET systems. As more operational data on individual components becomes available, this framework offers a natural space for refinement. Thus, the accuracy of the presented models is not an objective for this paper and the focus is on showing the possibility of training models on different variants of a system through using their subsystem behavior as input.

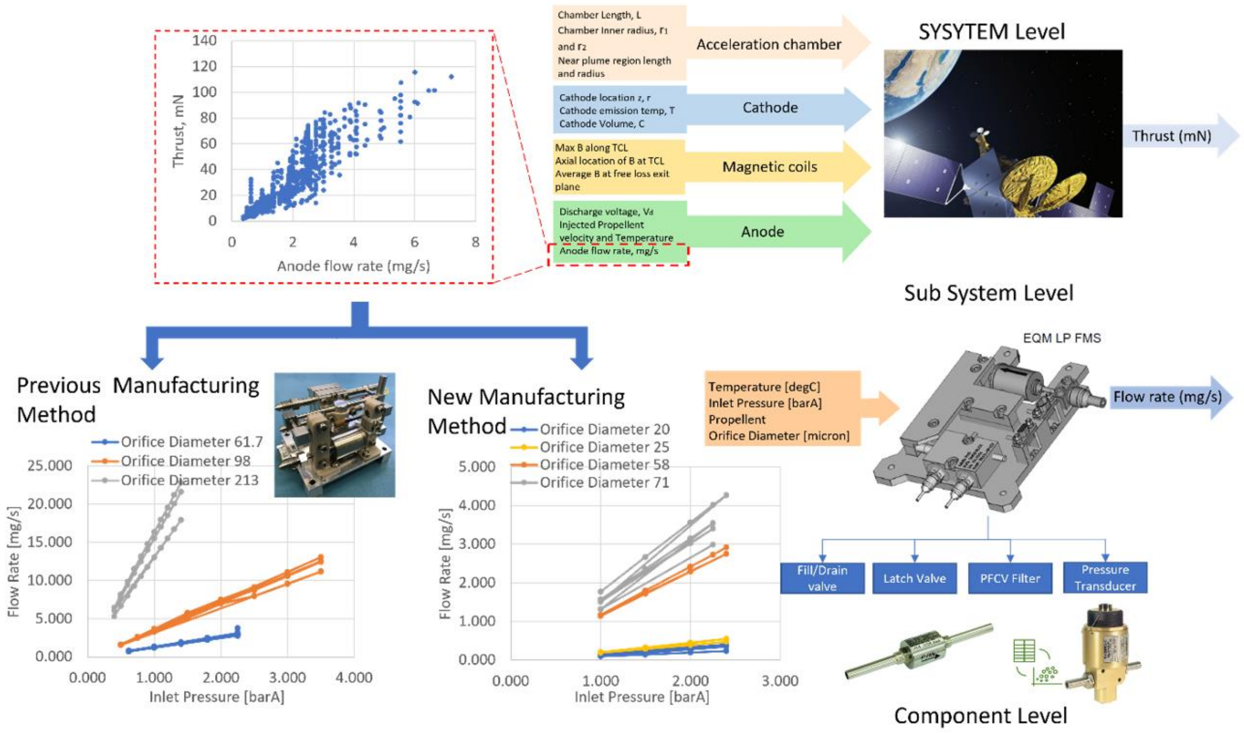

Figure 3. Performance of thruster and FMS defined by Flow rate as input in system level and output in subsystem level, Graphical elements representing system are derived from (Kuiper 2023)

4. Results

Building a global system-level ensemble envisioned in this paper requires building local models. To build multiple surrogate models that can be used to predict the performance of a new design case, a dataset of different hall effect thrusters was collected. HETs gathered in this paper represent a wide range of designs that have been proposed by different researchers over a long period of HET development. Therefore, the dataset encompasses different technologies with different operating conditions gathered from literature in this area. Used references for the dataset are not provided in this paper because of the scope of this study. Figure 4 shows collected 32 thruster results within different operating conditions that are tested with power between 100-2000 (W). Each dataset alone has somewhere from 5 to 20 test results that show a linear increase in thrust with the increase of anode flow rate and power. However, compiling all datasets together reveals an interesting nonlinear behavior that can be of interest to designers.

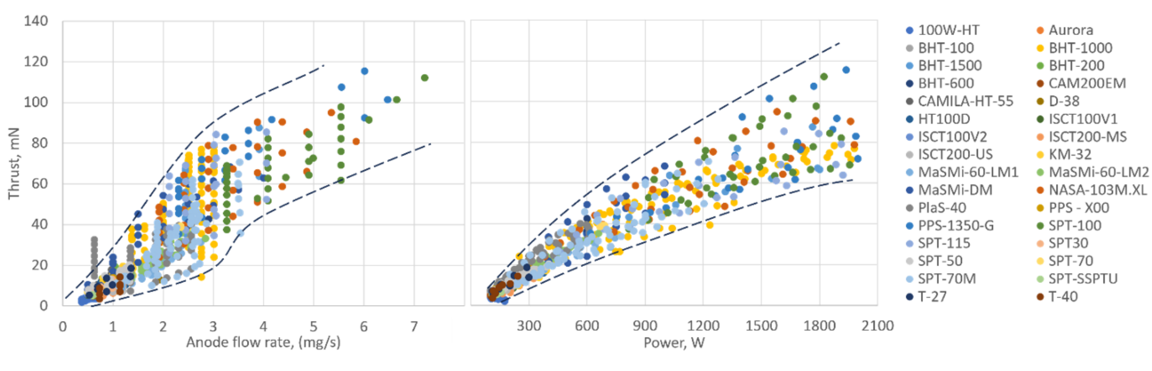

Figure 4. Thrust over Anode flow rate and Power for 32 thrusters from literature

In total, 32 surrogate models were trained per HET in the dataset. Each time, all other HETs are considered training sets and only one of the HETs is considered a testing set, in this way, the system behavior of HETs is captured in each model to make it possible to predict the performance of a new design case. For modeling purposes, after trying different algorithms, the Gaussian Process Regressor from scikit-learn library is used which is an open-source machine learning module in Python. The considered output as shown in the figure was "Thrust [N]" of the system, and for the input "Anode Flow Rate [mg/s]", "Anode Discharge Voltage [V]", and "Anode Discharge Current [A]" were used. In some cases where the reference paper did not provide the discharge power ("Discharge Power [W]"), the nominal power is used by multiplying the discharge current by the discharge voltage. After the training loop, trained models are compared in the predicted values and actual values that existed for each one of the HETs, this is annotated by the name of the thruster that it is predicting. Figure 5 shows the error rate between predicted and actual thrust in terms of Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2) for the thrusters in the dataset. The system-level prediction capability presented in Figure 5 shows an average RMSE of 2.85, an average MAPE of 11.25%, and an average R^2 of 0.77 for all 32 thruster data sets. It was also observed that the richer datasets performed better in terms of all three error matrices which is expected. This also explains the underperformance of those three HETs, which is clear from the figure.

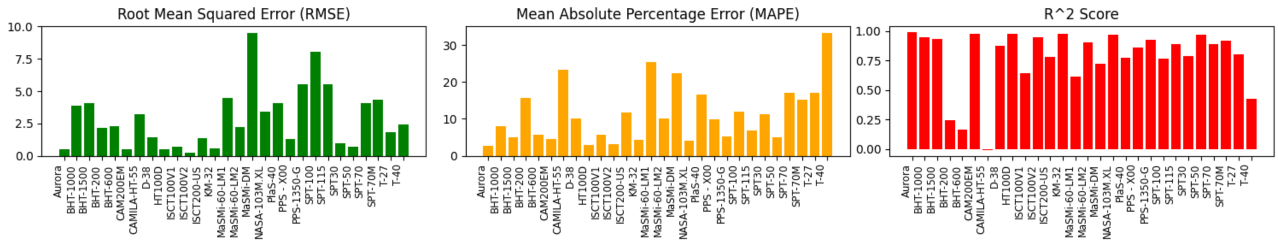

Figure 5. The error rate between predicted and actual for 32 thrusters for models trained

As mentioned, the HETs that exist in the dataset are of different studies, and training models on their performances using their operational condition is like extrapolating (widening) the boundaries of the design space. This data analysis enables a system perspective of what can be feasible to achieve from a future design. Overall, it can be inferred that using one group at a time as a testing set makes it possible to predict future HET designs based on operating conditions with a good error margin. Using this analysis domain designers can scale the product in the design space using the operational conditions that are used as input, however, this analysis alone doesn't say much about what can be expected if the designer changes one smaller unit (component level change) in the product architecture. At this point, we have only created the high-level predictive model of the envisioned ensemble and to be able to show the whole we need to zoom into subsystem and component levels.

FMS (Subsystem-level) predictiveness

To create the envisioned ensemble between different levels of product architecture in a modular way, we need to create shared interfaces between higher and lower-level datasets. Testing results for two alternative FMS architectures have been provided in this research by the company partner. These architectures are referred to as "previous manufacturing method" and "new manufacturing method". The results which are presented in the lowest row of Figure 4 show that for a certain orifice diameter, the relationship between flow rate and inlet pressure follows a near-linear trajectory. This is in line with the equation of continuity shown in Equation 1 and is valid for any incompressible fluid.

$$\dot{m} = \rho \times A \times \vartheta \qquad (1)$$

where the $\dot{m}$, ρ, A, and ϑ denotes mass flow rate, fluid density, cross-sectional area, and fluid velocity, respectively in fluid dynamics. It can be inferred that for a constant cross-sectional area of a pipe, the mass flow rate is proportional to the density and velocity of the fluid. This makes the density of a fluid directly proportional to its pressure, provided the temperature of the fluid is constant. Therefore, for a constant cross-sectional area of a pipe, the mass flow rate is proportional to the pressure of the fluid. This can be verified from test results presented in the lowest row of Figure 3 as well as the fundamental relationship in fluid systems, such as pipelines, valves, and nozzles. We also can conclude that mass flow rate is the linking feature between the two datasets (shared interface) allowing the ensemble in the next step.

One of the conditions to acquire the envisioned ensemble on product architecture is to be able to change products and their datasets in a modular way. To test this idea in practice, two alternative solutions at the subsystem level are used for training a machine-learning model. The alternative product with old manufacturing in the functional decomposition is considered as training and the new one as a testing set. Comparing the datasets of these alternatives shows although both datasets follow the same physics law (Equation 1) and share the same behavior, the ranges of the value and the proportion of the metrics to each other are different.

A neural net with three layers of [1024, 256, 32] neurons, with an L2 penalty of 0.01, and the ReLu activation function was created. It was found that adding L2 regularization to neural net architecture increased the accuracy of such out-of-range data (Ng 2004). Regularization is a technique used to prevent overfitting by adding a penalty term to the loss function. This penalty term encourages the model to have smaller weights, which in turn can help prevent overfitting. Therefore, since training and testing are from two different datasets (alternatives in product architecture) it is important that the model can extrapolate effectively. As explained, the training process used datasets for the "previous manufacturing method" and the testing was carried out with the "new manufacturing method" each representing two alternative solutions that exist for the same function as presented in the previous section.

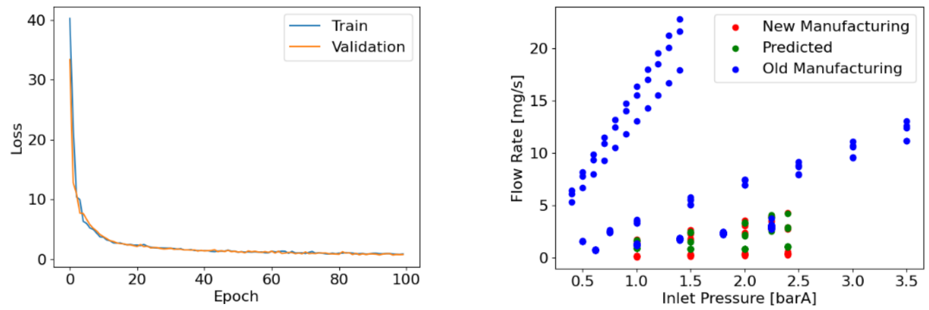

Figure 6. Training neural networks on old manufacturing methods and testing on new manufacturing method data

The results of the training in Figure 6 (left) show the loss function decreases monotonically throughout training in a smooth way, meaning that changes in the model's parameters result in small changes in the loss value. Training and validation loss curves converge to a similar and at a small value throughout the training which indicates that the model is fitting to the data in a good way. This figure (right) shows predicted and actual values for the validation dataset (New Manufacturing method). which shows the predictions were within an acceptable range. Moreover, it was found that the Mean Squared Error and R^2 error for the validation set were 0.21 and 0.86 respectively. That means it is possible to train a surrogate model on one product technology and predict the performance of another product technology based on shared functions in product architecture. This gives us another piece for building the envisioned ensemble.

Towards ensemble of surrogate models

One way to build an ensemble is to make use of the modeling capabilities of the Keras module (Chollet 2015) in Python to perform different operations over the layers/models. Stacking models on top of each other allows to combine the outputs of two or more models into a single output. For example, one can stack two convolutional neural networks (CNNs) to create a deeper CNN or stack a CNN with a recurrent neural network (RNN) to create a model that can process both spatial and temporal data. Another way to combine Keras models is to concatenate them over an axis as Keras models are essentially tensors. This is typically done when one wants to combine the inputs of two or more models into a single input. For example, two CNNs can be concatenated to create a wider CNN, or one CNN with an RNN can be concatenated to create a model that can process both spatial and temporal data in a parallel fashion. Additionally, the functional API in Keras which is one of the ways to create models, provides a more flexible way to combine models. This API ability allows the creation of models with arbitrary connectivity between layers and also makes it easier to reuse parts of models in different contexts.

Assuming a main model and a sub-model with their input and output layers. One can add the submodel output as an additional input to the main model and concatenate the main inputs and sub-output into a single input with tf.concat([main_input, sub_output], axis=1) shown in Figure 7. This figure also demonstrates how functionally decomposed product architecture in the case product can be broken down into thruster, FMS, and Valve level performances.

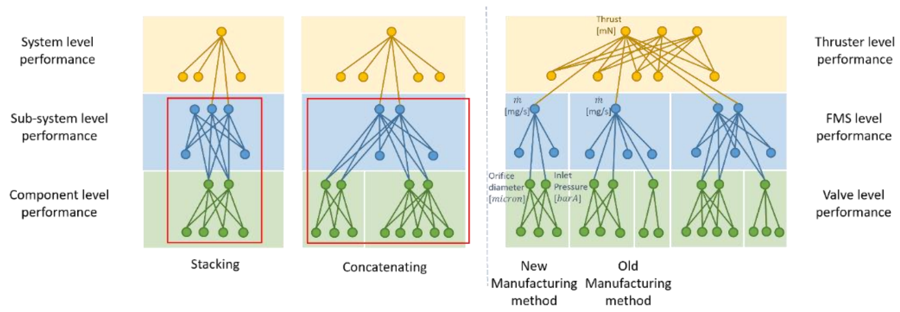

Figure 7. Stacking and Concatenating in Keras API and visualization of the ensemble on the HET case

5. Discussion and future work

Hall Effect Thrusters (HETs) are an ideal context to validate and expand a methodological framework integrating data-driven methods into early-stage design. This is because of their complexity, documented experimental data, well-defined subsystems, variations in design implementations, and broad operational and design spaces that offer a rich arena to test the robustness and extendibility of machine learning models. Successfully handling the multi-faceted design considerations of HETs points towards the promise of applying such a framework more broadly to other complex engineering systems.

Complex systems like HETs are usually tested physically or digitally to see if they pass requirements that were imposed on them during the design process. However, if we look broader into the performance of several designed systems, the observed behavior for different alternative systems is usually different from their constituting elements. Being able to predict the performance of HETs early in early design phases is important for selecting the right technologies for different functions. This is because, unlike other modeling and simulation fields, the development of these systems is highly dependent on physical testing and validation due to the high error rates of analytical methods and lack of accurate simulation models. This can be one of the reasons for the existence of scaling research filed in the HET literature that aims to predict the performance of an HET based on operating conditions, chamber geometry, and a series of other parameters. Most of the scaling literature is based on physics and analytical modeling (Andrenucci, Battista, and Piliero 2005; De Marco, Misuri, and Andrenucci 2007; Dannenmayer and Mazouffre 2011) while few recent are considering data-driven methods based on historical testing that are reported in the literature (Shagayda 2015; Plyashkov et al. 2022). Our envisioned ensemble can be positioned in a later group because it is leveraging data analysis on elements of a system to make global system predictions.

It has been shown that it is possible to change the dataset of a product over a common input-output (interface) to scale a solution space using functional decomposition. As shown in Figure 6, the Orifice diameter and Inlet pressure as two input features for the lower-level dataset enable us to have predictions over the Flow rate. The prediction error for Mean Squared Error of 0.21 and R^2 of 0.86 shows great generalizability for the ML models outside their trained space. This ability to train a model on one variant of a product and use the new product's input features to predict future variant results empowers domain engineers to foresee the results of their actions during the design phase. Moreover, since these two performed analyses are connected through a common feature (flow rate) as shown on the provided architecture it is possible to some extent to predict the global system output (thrust) with flow rate. The proposed avenue to build an ensemble can provide a way to connect two analyses. Using function-mean trees to find common features between different levels of analyses can provide a systematic solution for creating ensemble models that take in different methods and datasets which can be applied to different case products. This paper shows the characteristics and limitations of this method, and the result of the implementation remains for future studies.

The presented framework can provide the ability for "what-if" analysis regarding selecting different technological alternatives. The capacity to alter parameters and readily evaluate how system-level behavior responds mirrors one of the primary benefits of a digital twin environment. Even in the preliminary predictive mode, these capabilities enhance the understanding of the design space and promote proactive decision-making processes. As demonstrated in the previous section, this ensemble approach empowers designers to assess the system-wide impact of technology choices for individual components. Aligning product architecture, decomposed by functionality, with data-driven machine learning (ML) models offers a powerful framework for enhanced design space exploration. ML models, trained on input-output relationships at various levels of the product hierarchy, can bridge the gap between abstract functional requirements and predicted system performance. A shared interface between component-level and system-level models creates a flexible environment for rapid technology evaluation and scaling design solutions.

Future work for this study can be pursued at a detailed and process level. At the detail level, we aim to increase accuracy and achieve more robust models, recent regularization techniques such as PINNs can be incorporated into the equation of continuity in the ML models' loss function. In the process level and to extend the framework further development is necessary to merge the concepts of functional modeling and surrogate ensembles into a new way of predicting system-level performance. In this paper, we have presented a set of ideas for how this can be achieved.

6. Conclusions

This study demonstrates the potential for integrating machine learning (ML) in data analysis into early-phase design practices of complex engineered systems. Specifically, the framework was demonstrated over Hall Effect Thrusters (HETs) as a case study. A Gaussian Process Regression model trained over early HET test results (gathered from literature) effectively predicts HET thrust performance based on operational parameters like power and flow rate. This capability streamlines early-stage system-level insight into how early design cases can be useful in predicting future designs. A neural network was successfully used with two variations of subsystem design (Flow Management System) in HET. This showcases the framework's modularity and potential for technology assessment by predicting changes in performance stemming from component design or integration of new technologies. By suggesting an ensemble surrogate for design space exploration, the study lays the groundwork for a holistic system-level prediction environment. This integration, while requiring further development with techniques like Keras modeling, presents opportunities for improved decision-making. Designers can gain data-driven insights to understand trade-offs during conceptual design exploration, optimizing technology selection, and enhancing integration strategies.

While promising, this work also highlights essential directions for future exploration. Enhanced Dataset Collection requires more attention in the future. Creating large, publicly accessible datasets containing detailed operating and experimental data is crucial. This enables higher-performing ML models and promotes a collaborative research environment within the HET design field. Automated Feature Engineering can help develop a systematic approach for feature engineering that aligns domain knowledge with product architecture. Exploring techniques like Physics-Informed Neural Networks (PINNs) can offer opportunities to directly embed known physical laws into ML models, improving accuracy and reducing dependency on extensive testing data. Comprehensive Tool Implementation: Integrating the presented concepts into a software tool accessible to domain designers would greatly increase adoption and transform the design practice by accelerating design evaluation and iteration speeds.

Acknowledgment

This work has been supported by the CHEOPS LOW POWER project under Grant Agreement number: 101004331

References

Alizadeh, Reza, Liangyue Jia, Anand Balu Nellippallil, Guoxin Wang, Jia Hao, Janet K. Allen, and Farrokh Mistree. 2019. "Ensemble of Surrogates and Cross-Validation for Rapid and Accurate Predictions Using Small Data Sets." *AI EDAM* 33 (4): 484–501. https://doi.org/10.1017/S089006041900026X.

Andrenucci, Mariano, Francesco Battista, and Pietro Piliero. 2005. "Hall Thruster Scaling Methodology." In *The 29th Int. Electric Prop. Conf., Princeton University, 31 October–4 November 2005*.

Arjomandi Rad, Mohammad, Kent Salomonsson, Mirza Cenanovic, Henrik Balague, Dag Raudberget, and Roland Stolt. 2022. "Correlation-Based Feature Extraction from Computer-Aided Design, Case Study on Curtain Airbags Design." *Computers in Industry* 138 (June): 103634. https://doi.org/10.1016/j.compind.2022.103634.

Barai, S. V., and Yoram Reich. 1999. "Ensemble Modelling or Selecting the Best Model: Many Could Be Better than One." *AI EDAM* 13 (5): 377–86. https://doi.org/10.1017/S0890060499135029.

Boeuf, Jean-Pierre. 2017. "Tutorial: Physics and Modeling of Hall Thrusters." *Journal of Applied Physics* 121 (January): 011101. https://doi.org/10.1063/1.4972269.

Chen, Liming, Haobo Qiu, Chen Jiang, Xiwen Cai, and Liang Gao. 2018. "Ensemble of Surrogates with Hybrid Method Using Global and Local Measures for Engineering Design." *Structural and Multidisciplinary Optimization* 57 (4): 1711–29. https://doi.org/10.1007/s00158-017-1841-y.

Chollet, François. 2015. "Keras." https://keras.io.

Dannenmayer, Käthe, and Stéphane Mazouffre. 2011. "Elementary Scaling Relations for Hall Effect Thrusters." *Journal of Propulsion and Power* 27 (1): 236–45.

De Marco, Enrico Alessio, Tommaso Misuri, and Mariano Andrenucci. 2007. "A Review of the Hall Thruster Scaling Methodology."

Di Cara, Davina, and José Gonzales Del Amo. 2019. "EUROPEAN SPACE AGENCY, ESTEC, NOORDWIJK, THE NETHERLANDS." *InsideGNSS*, 2019.

Domínguez-Vázquez, Adrian, Jiewei Zhou, Alejandro Sevillano-González, Pablo Fajardo, and Eduardo Ahedo. 2022. "Analysis of the Electron Downstream Boundary Conditions in a 2D Hybrid Code for Hall Thrusters." In *37th International Electric Propulsion Conference, Electric Rocket Propulsion Society, Boston, MA*. https://ep2.uc3m.es/assets/docs/pubs/conference_proceedings/domi22a.pdf.

Fazeli, Hamid Reza, and Qingjin Peng. 2022. "Generation and Evaluation of Product Concepts by Integrating Extended Axiomatic Design, Quality Function Deployment and Design Structure Matrix." *Advanced Engineering Informatics* 54 (October): 101716. https://doi.org/10.1016/j.aei.2022.101716.

Geng, Xiuli, Xuening Chu, and Zaifang Zhang. 2010. "A New Integrated Design Concept Evaluation Approach Based on Vague Sets." *Expert Systems with Applications* 37 (9): 6629–38. https://doi.org/10.1016/j.eswa.2010.03.058.

Goel, Tushar, Raphael T. Haftka, Wei Shyy, and Nestor V. Queipo. 2007. "Ensemble of Surrogates." *Structural and Multidisciplinary Optimization* 33 (3): 199–216. https://doi.org/10.1007/s00158-006-0051-9.

Guo, Jiapeng, Ning Zhao, Lin Sun, and Saipeng Zhang. 2019. "Modular Based Flexible Digital Twin for Factory Design." *Journal of Ambient Intelligence and Humanized Computing* 10 (3): 1189–1200. https://doi.org/10.1007/s12652-018-0953-6.

Horizon. 2019. "Strategic Research Cluster – In-Space Electrical Propulsion and Station Keeping." Guidance document Version 1.0. https://ec.europa.eu/research/participants/data/ref/h2020/other/guides_for_applicants/h2020-supp-info-space-28-18-20_en.pdf.

Iansiti, Marco. 1995. "Technology Integration: Managing Technological Evolution in a Complex Environment." *Research Policy* 24 (4): 521–42. https://doi.org/10.1016/S0048-7333(94)00781-0.

Jiang, Zheng, Haobo Qiu, Ming Zhao, Shizhan Zhang, and Liang Gao. 2015. "Analytical Target Cascading Using Ensemble of Surrogates for Engineering Design Problems." *Engineering Computations* 32 (7): 2046–66. https://doi.org/10.1108/EC-11-2014-0242.

Khadke, Kiran, and John K. Gershenson. 2009. "Technology Change Analysis for Product and Product Platform Design." In , 461–70. American Society of Mechanical Engineers Digital Collection. https://doi.org/10.1115/DETC2007-35121.

Kriebel, Mary M. 2002. “System Engineering, Design, Integration, and Qualification of Electric Propulsion Space Experiment.” *Journal of Propulsion and Power* 18 (4): 731–39. https://doi.org/10.2514/2.6020.

Kuiper, Johan. 2023. “Initial High Pressure Krypt Functional Performance Results of the Fluid Management System.” EPIC WORKSHOP. Naples: BRADFORD Space. https://www.epic-src.eu/wp-content/uploads/8-Johan-Kuiper_BRADFORD.pdf.

Lai, Xiaonan, Xiwang He, Yong Pang, Fan Zhang, Dongcai Zhou, Wei Sun, and Xueguan Song. 2023. “A Scalable Digital Twin Framework Based on a Novel Adaptive Ensemble Surrogate Model.” *Journal of Mechanical Design* 145 (2): 021701.

Liu, Mengnan, Shuiliang Fang, Huiyue Dong, and Cunzhi Xu. 2021. “Review of Digital Twin about Concepts, Technologies, and Industrial Applications.” *Journal of Manufacturing Systems* 58: 346–61.

Mienye, Ibomoiye Domor, and Yanxia Sun. 2022. “A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects.” *IEEE Access* 10: 99129–49. https://doi.org/10.1109/ACCESS.2022.3207287.

Müller, Jakob R., Ola Isaksson, Jonas Landahl, Visakha Raja, Massimo Panarotto, Christoffer Levandowski, and Dag Raudberget. 2019. “Enhanced Function-Means Modeling Supporting Design Space Exploration.” *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 33 (4): 502–16. https://doi.org/10.1017/S0890060419000271.

Ng, Andrew Y. 2004. “Feature Selection, L1 vs. L2 Regularization, and Rotational Invariance.” In *Proceedings of the Twenty-First International Conference on Machine Learning*, 78. ICML ’04. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/1015330.1015435.

Ogunsakin, Rotimi, Nikolay Mehandjiev, and Cesar A. Marin. 2023. “Towards Adaptive Digital Twins Architecture.” *Computers in Industry* 149 (August): 103920. https://doi.org/10.1016/j.compind.2023.103920.

Ölvander, Johan, Björn Lundén, and Hampus Gavel. 2009. “A Computerized Optimization Framework for the Morphological Matrix Applied to Aircraft Conceptual Design.” *Computer-Aided Design*, Computer Support for Conceptual Design, 41 (3): 187–96. https://doi.org/10.1016/j.cad.2008.06.005.

Panarotto, Massimo, Ola Isaksson, and Leif Asp. 2018. “ASSESSING THE VALUE OF RADICAL TECHNOLOGY ALTERNATIVES AT SYSTEM LEVEL.” In *DS 92: Proceedings of the DESIGN 2018 15th International Design Conference*, 633–42. https://doi.org/10.21278/idc.2018.0398.

Plyashkov, Yegor V., Andrey A. Shagayda, Dmitrii A. Kravchenko, Alexander S. Lovtsov, and Fedor D. Ratnikov. 2022. “On Scaling of Hall-Effect Thrusters Using Neural Nets.” *Journal of Propulsion and Power* 38 (6): 935–44. https://doi.org/10.2514/1.B38592.

Qi, Jin, Jie Hu, and Yinghong Peng. 2021. “A Customer-Involved Design Concept Evaluation Based on Multi-Criteria Decision-Making Fusing with Preference and Design Values.” *Advanced Engineering Informatics* 50 (October): 101373. https://doi.org/10.1016/j.aei.2021.101373.

Qi, Qinglin, Fei Tao, Tianliang Hu, Nabil Anwer, Ang Liu, Yongli Wei, Lihui Wang, and A. Y. C. Nee. 2021. “Enabling Technologies and Tools for Digital Twin.” *Journal of Manufacturing Systems*, Digital Twin towards Smart Manufacturing and Industry 4.0, 58 (January): 3–21. https://doi.org/10.1016/j.jmsy.2019.10.001.

Schachinger, Peter, and Hans L. Johannesson. 2000. “Computer Modelling of Design Specifications.” *Journal of Engineering Design* 11 (4): 317–29. https://doi.org/10.1080/0954482001000935.

Scime, Luke, Alka Singh, and Vincent Paquit. 2022. “A Scalable Digital Platform for the Use of Digital Twins in Additive Manufacturing.” *Manufacturing Letters* 31: 28–32.

Shagayda, Andrey A. 2015. “On Scaling of Hall Effect Thrusters.” *IEEE Transactions on Plasma Science* 43 (1): 12–28. https://doi.org/10.1109/TPS.2014.2315851.

Snyder, John Steven, Jeff Baldwin, Jason D. Frieman, Mitchell LR Walker, Nathan S. Hicks, Kurt A. Polzin, and James T. Singleton. 2013. “Flow Control and Measurement in Electric Propulsion Systems: Towards an Aiaa Reference Standard.” In *33rd International Electric Propulsion Conference*, 2013–2425. Electric Rocket Propulsion Soc. Fairview Park, OH. https://apps.dtic.mil/sti/citations/ADA591932.

Stork, Jörg, Martina Friese, Martin Zaefferer, Thomas Bartz-Beielstein, Andreas Fischbach, Beate Breiderhoff, Boris Naujoks, and Tea Tušar. 2020. “Open Issues in Surrogate-Assisted Optimization.” In *High-Performance Simulation-Based Optimization*, edited by Thomas Bartz-Beielstein, Bogdan Filipič, Peter Korošec, and El-Ghazali Talbi, 833:225–44. Studies in Computational Intelligence. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-18764-4_10.

Tao, Fei, Fangyuan Sui, Ang Liu, Qinglin Qi, Meng Zhang, Boyang Song, Zirong Guo, Stephen C.-Y. Lu, and A. Y. C. Nee. 2019. “Digital Twin-Driven Product Design Framework.” *International Journal of Production Research* 57 (12): 3935–53. https://doi.org/10.1080/00207543.2018.1443229.

Tao, Fei, Bin Xiao, Qinglin Qi, Jiangfeng Cheng, and Ping Ji. 2022. “Digital Twin Modeling.” *Journal of Manufacturing Systems* 64 (July): 372–89. https://doi.org/10.1016/j.jmsy.2022.06.015.

Ulrich, Karl T., and Steven D. Eppinger. 2016. *Product Design and Development*. McGraw-hill. https://thuvienso.hoasen.edu.vn/handle/123456789/9147.

Ulrich, Karl T., Steven D. Eppinger, and Maria C. Yang. 2008. *Product Design and Development*. Vol. 4. McGraw-Hill higher education Boston.

Vial, Vanessa. 2023. “Consortium for Hall Effect Orbital Propulsion System.” EPIC WORKSHOP. Naples. https://www.epic-src.eu/wp-content/uploads/9-EPIC-WS-CHEOPS-presentation-2023-05-10-zz.pdf.

Wang, G. Gary, and S. Shan. 2006. “Review of Metamodeling Techniques in Support of Engineering Design Optimization.” *Journal of Mechanical Design* 129 (4): 370–80. https://doi.org/10.1115/1.2429697.

Yuan, Chenxi, Tucker Marion, and Mohsen Moghaddam. 2021. “Leveraging End-User Data for Enhanced Design Concept Evaluation: A Multimodal Deep Regression Model.” *Journal of Mechanical Design* 144 (021403). https://doi.org/10.1115/1.4052366.

Zwicky, Fritz. 1969. “Discovery, Invention, Research through the Morphological Approach.” https://philpapers.org/rec/ZWIDIR.

Contact: Mohammad Arjomandi Rad, Chalmers University of Technology, Department of Industrial and Material Science, Hörsalsvägen 7A, 412 58, Gothenburg, Sweden, radmo@chalmers.se