ML4LM: Machine Learning for Safely Landing on Mars

David D. Wu*
Stanford University
450 Serra Mall
Stanford, CA 94305
daaw@stanford.edu

Wai Tong Chung†
Stanford University
450 Serra Mall
Stanford, CA 94305
wtchung@stanford.edu

Matthias Ihme‡
Stanford University
450 Serra Mall
Stanford, CA 94305
SLAC National Accelerator Laboratory
2575 Sand Hill Rd
Menlo Park, CA 94025
mihme@stanford.edu

Abstract

Human missions to Mars will require rocket-powered descent through the Martian atmosphere to safely land. Designing the propulsion system for these missions introduces the following challenges: ensuring safety, sparsity of data to validate models, and requirements for rapid predictions. ML offers opportunities for addressing these challenges. We use ML methods to develop novel data-analytic tools that support design analysis for enabling supersonic retropropulsion (SRP) technology deployment. Accordingly, we propose a hierarchical physics-embedded data-driven (HPDD) framework for predicting the key target quantity in SRP. HPDD model is trained on experimental data from wind tunnel testing (small-scale rocket), and the model exhibits promising accuracy and computational efficiency. Wind tunnel testing in the future will provide more data for validation and enhancement of our framework to further the understanding of SRP.

1 Introduction

Supersonic retropropulsion (SRP) is a type of powered descent that uses rocket propulsion against freestream flow to decelerate a spacecraft. SRP is a critical technology for safely landing human-scale vehicles on Mars [3, 9]. Figure 1 shows a Schlieren wind tunnel image of an SRP rocket exhaust plume with freestream flow flowing from left to right [2].

Modeling SRP is challenging because safety is paramount, the available data is sparse, and conventional high-fidelity modeling approaches can be prohibitively expensive (especially for engineering design processes). To account for these challenges, our key objectives are to propose data-analytic methods to develop models for predicting multi-nozzle plume physics, and analyze sensitivities for enabling SRP technology.

Physics-informed ML is uniquely capable of addressing these challenges beyond traditional methods' (analytical reduced-order models, computational fluid dynamics) capabilities in both modeling accuracy and computational efficiency. We combine physical knowledge with the relatively small wind tunnel dataset to ensure accuracy. Moreover, we believe that our modeling framework will eventually aid in exploring novel rocket designs for faster engineering design.

*Graduate Student.

†Graduate Student. Human-Centered Artificial Intelligence Graduate Fellow.

‡Professor of Mechanical Engineering and of Photon Science.

Machine Learning and the Physical Sciences workshop, NeurIPS 2022.

Figure 1: Schlieren image data of rocket (horizontal black rectangle) with visible exhaust plume thrusting against supersonic freestream flow. The bow shock is the curved edge that spans figure from top to bottom.

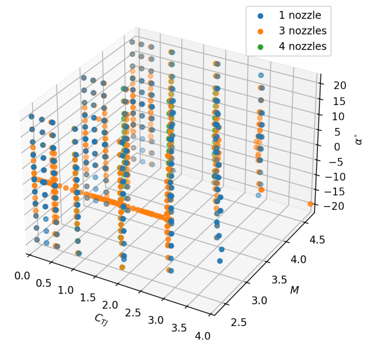


Figure 2: Input feature space of our data set for all nozzle configurations, showing sparsity. Axes: total thrust coefficient of rocket jet C_{TJ}, Mach number M, and angle of attack in degrees α°.

ML allows us to effectively extract insights from the data to build an accurate and inexpensive model. To reduce gaps in both understanding and data, we decided on three key objectives. First, we develop and apply data-analysis methods to extract features that capture plume physics from wind tunnel data to enable SRP deployment. Second, using this knowledge, we develop a hierarchical physics-embedded data-driven (HPDD) framework for the key target quantity C (coefficient of aerodynamic axial force) that captures plume physics, plume-aerodynamics interaction, and sensitivities. Third, we reduce uncertainty by identifying data gaps to be filled by future wind tunnel tests. Overall, all three objectives strive toward improved modeling capabilities of complicated plume structures and associated physics, ultimately for enabling human Mars missions.

2 Method

2.1 Data

We use a dataset [2] from the NASA Langley Unitary Plan Wind Tunnel, consisting of 2500 data points. The data has 168 features θ (67 independent features plus 101 engineered features), and 125 measurements. According to our physics intuition, the following are key features that span flight-relevant spaces: Mach number M, total thrust coefficient of rocket jet C_{TJ}, nozzle configuration $c_{\phi qe}$ (calculated by exponential of Morris-Mitchell Criterion [8]), angle of attack α, sideslip angle β, freestream stagnation pressure $P_{T\infty}$, freestream static pressure P_{∞}, freestream static temperature T_{∞}, freestream post-normal shock stagnation pressure P_{T2}, and freestream Reynolds number Re_{∞}. The output measurements are pressure data from the pressure probes on the rocket forebody. Overall, this dataset is high-dimensional with relatively few data points.

To pre-process this data, we use Lagrange interpolation with Delaunay triangulation to compute C from the measurements. This maps the 125-dimensional measurements vector to a scalar. Figure 2 shows three features relevant to rocket propulsion; the data is sparse across C_{TJ} and M. Due to data limitations, we focus on the supersonic regime; if given more comprehensive data, we could examine model extrapolation ability to other regimes.

2.2 Modeling framework

Using the data described above, we aspire to develop a HPDD model that accurately and cost-efficiently predicts C over a range of flight-relevant operating conditions. We seek to address the following limitations of solely relying on neural networks (NN) for regression. First, because of their low interpretability, it can be difficult to anticipate their outputs. We are working with a safety-critical application, where failure can cause loss of human life, therefore we require well-behaved outputs.

Second, NNs are not necessarily able to generalize to problems that are outside of the given data space, especially given the limited scope of our experimental dataset.

We address the above limitations by first leveraging physical knowledge to both reduce our dependence on the NN, and enhance our NN design. Physics shows that for a given thrust regime, the magnitudes of the SRP targets are related [10]; therefore, we use one fully connected NN to regress residuals, rather than using one NN for each target. Physical knowledge also enhances the model's robustness and generalization capabilities for unexplored rocket designs.

To accomplish our key objectives, we introduce a hierarchical physics-embedded data-driven framework which combines ML with physics knowledge. Existing work for relevant applications [9, 1] provides physical interpretation, while the available data improves accuracy (and thus safety) by compensating for limitations introduced by the physics-based model. This hybrid approach (closely resembling [6]) also requires less data than purely data-driven approaches. With this, our data-driven framework F can be written as:

$$F(\theta_p, \theta_c, \theta_u) = \mathcal{P}(\theta_p) * \mathcal{C}(\theta_c) + \mathcal{U}(\theta_u), \tag{1}$$

where $\mathcal{P}$ is a physics-embedded reduced-order model that describes plume-interaction and coalescence [9, 1]. $\mathcal{C}$ is a coupling function that uses a NN to model non-linear phenomena and is based on data-driven algorithms. $\mathcal{U}$ is an uncertainty quantification model to improve interpretability of F. For the inputs, we desire $\theta_p \subset \theta_c \subset \theta$ construction; section 2.3 will show how we use both data analytics and physics knowledge to robustly identify the important features of the SRP flow field.

We want to compare cases where the rocket is off/on. To this end, we convert Equation (1) into an additive HPDD model for C:

$$C = \underbrace{C_0 + C_{SRP}}_{\mathcal{P}(\theta_p)} + \underbrace{C_{NN}}_{\mathcal{C}(\theta_c)} \tag{2}$$

The additive form of Equation (2) allows us to examine the contribution of each term, and is useful for gaining insight into how turning the rocket on affects C. C is the sum of a zero thrust vehicle model C_0 (rocket off), a model with SRP C_{SRP} (rocket on), and a multilayer perceptron residual regression model C_{NN}.

2.3 Identify input feature sets

We apply both unsupervised and supervised data-driven feature selection algorithms to inform our choice of a salient input set θ_p to characterize $\mathcal{P}$. The algorithms are: principal component analysis (PCA) with 3 principal components to identify correlations in the input space, correlation heatmap to quantify input-output relationship strength, and random forest (RF) with 200 trees to quantify input importance. Figure 3 shows our heatmap result, and figure 4 shows our RF regressor result with target C. Physical intuition in addition to both PCA and the heatmap emphasize the importance of M. Both the heatmap and RF emphasize the importance of C_{TJ} and $c_{\phi qe}$. Our physical intuition informs us that α is important. Therefore, we determine $\theta_p = \{M, C_{TJ}, c_{\phi qe}, \alpha\}$.

To find the input set θ_c for C_{NN}, we use physical knowledge [9, 1, 5, 7] and domain expertise to determine $\theta_c = \{M, C_{TJ}, c_{\phi qe}, \alpha, \beta, P_{T\infty}, P_\infty, T_\infty, P_{T2}, Re_\infty\}$.

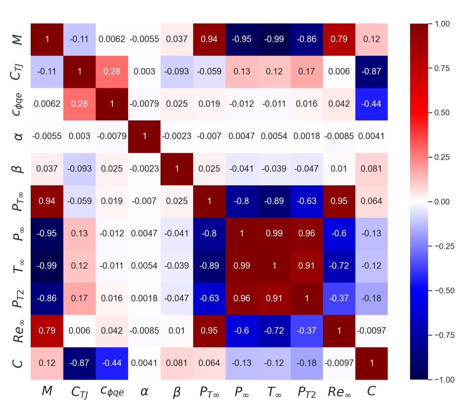


Figure 3: Heatmap relationship strength.

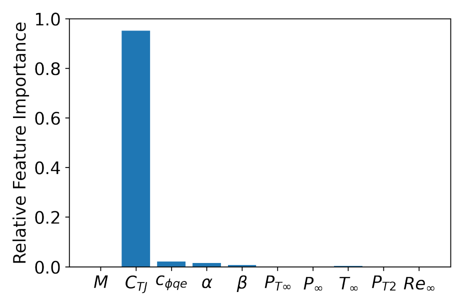


Figure 4: Feature importance using RF. C_{TJ}, and $c_{\phi qe}$ are the most important features. This result is consistent with heatmap result.

2.4 NN setup and computational cost

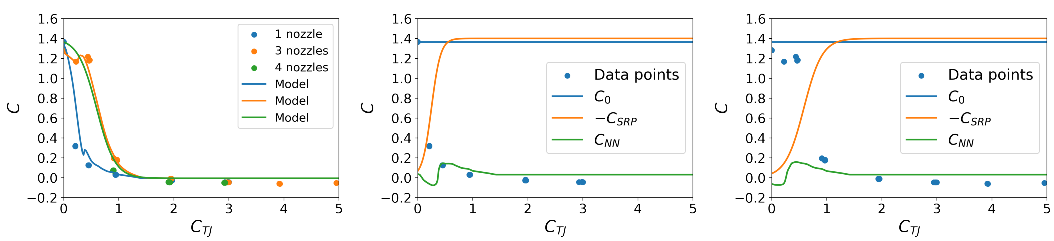
We use Keras [4] to train a feedforward NN residual model using as input θ_c, and as target C. Our normalization technique is linear scaling. We perform small scale hyperparameter search, architecture search, and algorithm search to determine the following NN architecture. We use four hidden dense layers where the number of neurons is 400, 200, 100, and 50; the NN has 109,954 total parameters. Each hidden layer uses ReLU activation, and we reduce overfitting through the use of 30% dropout. We optimize using Adam (learning rate $= 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-7}$) with respect to mean squared error loss. Our train/validation/test split is 80/10/10, and we train for 100 epochs with a batch size of 5 data points. We spend 105 seconds training, and we set the seed equal to 0 for reproducibility. Finally, our model spends on average 0.03 seconds per prediction (hardware: i7-9750H CPU, 2.6GHz), meaning there is potential for real-time vehicular deployment with dynamics and controls systems.

3 Results

Equation (3) presents our HPDD model for C:

$$C = \underbrace{\frac{4\tan(\delta)}{\gamma M^2}\left(\frac{2\gamma M^2 - (\gamma - 1)}{\gamma + 1} - 1\right)\cos^{1.3}(\alpha)}_{\text{Zero thrust vehicle model } C_0} + \underbrace{\frac{L}{1 + e^{-k(C_{TJ} - x_0)}}}_{\text{Model with SRP } C_{SRP}} + \underbrace{C_{NN}(\theta_c)}_{\text{NN residual}} \tag{3}$$

We account for curved shocks by fitting $\delta = \phi_r^\circ + 0.372M^2 - 3.467M - 38.422$, where ϕ_r is half-angle of rocket forebody. γ is ratio of specific heats, assumed constant value of 1.4. C_0 is based on functional forms from literature [1, 9]. C_{SRP} uses knowledge of plume structure transition physics [10] to inform our choice to use the logistic function as the basis function for fitting. We account for Mach effects using data to fit $L = -0.0124M^2 + 0.1186M - 1.663$. We model multi-nozzle physical effects using data to fit k and x_0 as follows: $k = e^{-3(c_{\phi qe} - 1.6)} + 6$, and $x_0 = -e^{-1.05c_{\phi qe}} + 0.6$. Finally, we add C_{NN} according to Equation (2).



(a) Various rocket nozzles. (b) 1 nozzle. (c) 3 nozzles.

Figure 5: Model plotted against test data points for $M = 3.5$, $\alpha = 0°$, and $\beta = 0°$.

Figure 5 plots HPDD model predictions against data points from the test set. Figures 5b and 5c show that the zero thrust vehicle model accurately fits the data points when $C_{TJ} = 0$. Moreover, the model with SRP decreases C as C_{TJ} increases because of expanding plume blockage. The NN residual regression provides small corrections, that account for non-linear coupling, in the low thrust regime where both turbulence and unsteadiness can be observed in the Schlieren image data. Figure 5b shows reasonable extrapolations (for $C_{TJ} > 3$) that are consistent with physical knowledge [9]. We are confident in the accuracy and interpretability of our NN because it consistently provides very small magnitude corrections in the flight-relevant regime (high thrust) [3]; high thrust inputs cause the NN to deactivate. Such a well-behaved result supports the safety-critical dimension of our application.

4 Conclusions

Overall, we propose a hierarchical physics-embedded data-driven framework that leverages limited data for enhancing safety and computational-efficiency for enabling SRP technology. This framework consists of a low-order physical model, an ML coupling function, and uncertainty quantification. $\mathcal{P}$

ensures accurate interpolations and extrapolations, while $\mathcal{C}$ exploits the ability of ML to fit non-linear, usually elusive, phenomenon (ex. rocket exhaust plume coalescence [7]). We use data-driven methods to verify inputs for the physics-embedded model, and we use physics domain knowledge to determine inputs for the ML model.

We believe the above results are promising and will extend our framework to other targets in the future. As more future work, we will validate our models with more data (i.e. 4 nozzles), employ rigorous uncertainty quantification, and potentially extend our modeling framework to other regimes (ex. transonic, subsonic). Ultimately, we hope to see its application on a human spacecraft to Mars in the not-too-distant future.

Broader impact statement

There would be at least two potential positive impacts generated by more accurate and affordable modeling of SRP. First, it can better inform the engineering design of human-scale Mars landers because our models' insights improve the understanding of SRP physics. That means our models can enable faster prototyping if they are determined to be appropriate for use in conjunction with prototype testing of landers. Second, the versatility of our modeling framework means that it has the potential to be adapted to solve related problems in rocket propulsion. In terms of potential negative societal impacts, model extrapolation can have unexpected consequences during prototype testing. If the model's output is inaccurate compared to tests, we can enhance the model; additionally, we can introduce uncertainty quantification to provide estimates on the uncertainty of the model.

Acknowledgments and Disclosure of Funding

We are grateful to Karl Edquist and Ashley Korzun for their domain expertise and many insightful discussions. Funding in direct support of this work: NASA grant 80NSSC22K0257, Early Stage Innovations Space Technology Research Grants Program.

Competing interests: none for all authors.

References

[1] J. D. Anderson. *Fundamentals of Aerodynamics*. McGraw-Hill Education, 6th edition, 2017.

[2] S. Berry, M. Rhode, K. Edquist, and C. Player. Supersonic retropropulsion experimental results from the NASA Langley Unitary Plan Wind Tunnel. *42nd AIAA Thermophysics Conference*, Jun 2011. doi: 10.2514/6.2011-3489.

[3] R. D. Braun, B. Sforzo, and C. Campbell. Advancing supersonic retropropulsion using mars-relevant flight data: An overview. *AIAA SPACE and Astronautics Forum and Exposition*, 2017. doi: 10.2514/6.2017-5292.

[4] F. Chollet et al. Keras. `https://keras.io`, 2015.

[5] E. O. Daso, V. E. Pritchett, T.-S. Wang, D. K. Ota, I. M. Blankson, and A. H. Auslender. Dynamics of shock dispersion and interactions in supersonic freestreams with counterflowing jets. *AIAA Journal*, 47(6):1313–1326, 2009. doi: 10.2514/1.30084.

[6] R. Delli Compagni, Z. Cheng, S. Russo, and T. P. Van Boeckel. A hybrid Neural Network-SEIR model for forecasting intensive care occupancy in Switzerland during COVID-19 epidemics. *PLOS ONE*, 17:1–15, 03 2022. doi: 10.1371/journal.pone.0263789.

[7] K. T. Edquist, S. J. Alter, C. E. Glass, W. L. Kleb, A. M. Korzun, W. A. Wood, F. Canabal, R. Childs, L. D. Halstrom, and K. V. Matsuno. Computational modeling of Mars retropropulsion concepts in the Langley Unitary Plan Wind Tunnel. *AIAA SCITECH 2022 Forum*, Jan 2022. doi: 10.2514/6.2022-0912.

[8] M. J. Kochenderfer and T. A. Wheeler. *Algorithms for Optimization*. The MIT Press, 2019.

[9] A. M. Korzun. *Aerodynamic and Performance Characterization of Supersonic Retropropulsion for Application to Planetary Entry and Descent*. PhD thesis, Georgia Institute of Technology, 2012.

[10] K. A. Montgomery, P. J. Bruce, and S. Navarro-Martinez. Dynamics of varying thrust coefficients for supersonic retropropulsion during Mars EDL. *AIAA SCITECH 2022 Forum*, Jan 2022. doi: 10.2514/6.2022-0935.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes]
   (c) Did you discuss any potential negative societal impacts of your work? [Yes]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [No]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [No]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [No]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]