Sample Efficient Modeling of Drag Coefficients for Satellites with Symmetry

Neel Sortur SORTUR.N@NORTHEASTERN.EDU
Linfeng Zhao ZHAO.LINF@NORTHEASTERN.EDU
Robin Walters R.WALTERS@NORTHEASTERN.EDU
Northeastern University

Editors: Sophia Sanborn, Christian Shewmake, Simone Azeglio, Nina Miolane

Abstract

Accurate knowledge of the atmospheric drag coefficient for a satellite in low Earth orbit is crucial to plan an orbit that avoids collisions with other spacecraft, but its calculation has high uncertainty and is very expensive to numerically compute for long-horizon predictions. Previous work has improved coefficient modeling speed with data-driven approaches, but these models do not utilize domain symmetry. This work investigates enforcing the invariance of atmospheric particle deflections off certain satellite geometries, resulting in higher sample efficiency and theoretically more robustness for data-driven methods. We train G-equivariant MLPs to predict the drag coefficient, where G defines invariances of the coefficient across different orientations of the satellite. We experiment on a synthetic dataset computed using the numerical Test Particle Monte Carlo (TPMC) method, where particles are fired at a satellite in the computational domain. We find that our method is more sample and computationally efficient than unconstrained baselines, which is significant because TPMC simulations are extremely computationally expensive.

1. Introduction

Earth's atmosphere induces a drag force on satellites in low Earth orbit (LEO). This drag needs to be modeled in order to accurately propagate orbits into the future to determine if any satellites are at risk of colliding, preventing debris cascades (Kessler et al., 2010). A key part of modeling this force is obtaining a drag coefficient, a parameter that captures the interaction between atmospheric particles and a satellite's surface. This parameter has closed-form solutions for simple satellite geometries like a sphere (Chambre and Schaaf, 1961), but more complex shapes need numerical simulations of atmosphere-satellite interaction. Because these simulations are too slow for the desired orbit propagation, previous work has used Bayesian and deep learning techniques (Paul et al., 2023) to learn a faster model of the dynamics system. However, these methods still rely on the simulations for vast quantities of training data, sometimes infeasible to produce with a reasonable amount of time and compute.

In this work, we find that modeling drag coefficients can be made more sample efficient by utilizing symmetry of the satellite geometry, since the satellite deflects atmospheric particles invariantly under certain orientations. This efficiency makes data-driven methods more feasible. Using the `escnn` framework (Cesa et al., 2022), we design equivariant neural networks invariant to the symmetry group of the standard CubeSat, a cube shaped satellite, and the Gravity Recovery and Climate Experiment (GRACE) satellite (Tapley et al., 2004). These networks also provide better generalization ability and accuracy than unconstrained versions (Wang et al., 2021).

2. Background

Drag Force and Coefficient Models. The force acting on a satellite in LEO is a function of satellite mass, velocity, cross sectional area, atmospheric mass density, and also the drag coefficient. We focus on modeling the drag coefficient C_d, a large source of uncertainty in this function. Its calculation is affected by the choice of gas-surface interaction (GSI) model, which has been previously studied (Mehta et al., 2014b). Here, we use the Cercignani–Lampis–Lord (CLL) model (Cercignani and Lampis, 1971). This model computes C_d in terms of the following inputs, which we will refer to as $\boldsymbol{x} \in \mathbb{R}^7$, where x_1 is relative velocity of the satellite, x_2 is satellite surface temperature, x_3 is atmospheric translational temperature, x_4 is normal energy accommodation coefficient, x_5 is tangential momentum accommodation coefficient, x_6 is satellite yaw, and x_7 is satellite pitch. Yaw and pitch together define the satellite's orientation, also known as the *attitude*.

Numerical Simulation of Drag Coefficient. The RSM Toolkit (Walker and Lawrence, 2016) uses the Test Particle Monte Carlo (TPMC) simulation technique to produce a drag coefficient $C_d \in \mathbb{R}$ for each $\boldsymbol{x}$. It works by sequentially firing a number of molecules of real gas into the computational domain, which deflect off a satellite surface specified by a stereolithography (STL) file. These particles do not undergo intermolecular collisions because LEO is in the Free Molecular Flow regime (Chambre and Schaaf, 1961).

Groups and Equivariant Networks. A symmetry group G is a set with an associative binary operation that has inverses and an identity. A *group representation*, denoted by $\rho(g)$, associates an $n \times n$ matrix with each $g \in G$ in a way which is compatible with the multiplication table of G, which we will use to apply changes to satellite attitude. Given a group G and representations ρ_X and ρ_Y acting on X and Y, we say a function $f\colon X \to Y$ is *equivariant* if, for all $x \in X$, $g \in G$, $\rho_Y(g) \cdot f(x) = f(\rho_X(g) \cdot x)$.

A standard MLP consists of linear layers $y = Mx$ and non-linear activations σ. A G-equivariant MLP is composed of linear layers $y = Mx$ in which the matrices M *intertwine* the group action $\rho_Y(g)M = M\rho_X(g)$ (Finzi et al., 2021). The space of such M is determined before training by solving a linear system. Training is then restricted to this class of matrices. The non-linearity σ must also be equivariant, but for permutation representations, standard pointwise nonlinearities are equivariant. We use *group pooling* to make y invariant to g such that $\texttt{Pooling}(y) = \texttt{Pooling}(\rho_{\texttt{out}}(g) \cdot y)$ for all $g \in G$. (Cohen and Welling, 2016).

To encode representations in these MLPs, let us consider $G_1, G_2 = C_2$ (the cyclic group of order 2) used for reflection. ρ_0 is the trivial representation (1×1 identity matrix), and ρ_1 is the sign representation (1×1 matrix for 1 or -1). We need to use representations of the direct product group $G_1 \times G_2$, which are defined by *exterior tensor products* of the representations of each individual group $\rho_i^{G_1} \boxtimes \rho_j^{G_2}$.

3. Method

3.1. Geometric Properties of TPMC

The CLL TPMC method uses the incident and reflected velocities of computational particles to determine the gas-surface interaction, and thus C_d. The incident velocity is fixed for all simulations. Let us define the reflected velocity as a function of the satellite geometry *mesh*,

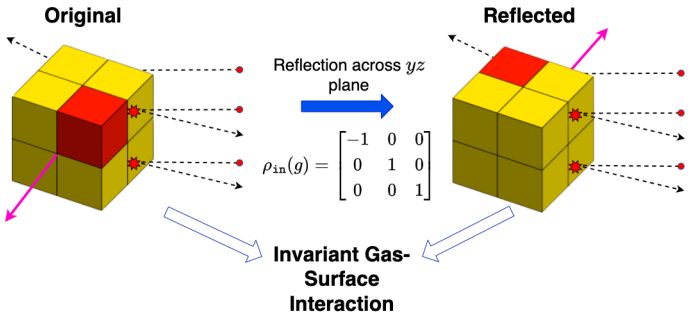

Figure 1: Invariance of GSI for cube shaped satellite under an arbitrary reflection across the yz plane. The red dots are atmospheric gas particles. The pink vector represents the heading, and the red cube differentiates the original geometry from the reflected geometry.

the variables $x_1, \ldots, x_5$ as the *physical* variables $\boldsymbol{x}_p \in \mathbb{R}^5$, and the variables x_6, x_7 as the *attitude* variables $\boldsymbol{x}_a \in \mathbb{R}^2$. Then $C_d = f(\boldsymbol{x}_p, \boldsymbol{x}_a | mesh)$.

Under the symmetry group G of the satellite geometry, the reflected velocity is invariant to actions of G acting on the satellite attitude: $\rho_{\texttt{out}}(g) \cdot f(\boldsymbol{x}_p, \boldsymbol{x}_a | mesh) = f(\boldsymbol{x}_p, \rho_{\texttt{in}}(g) \cdot \boldsymbol{x}_a | mesh)$. Consider $G = C_2 \times C_2 \times C_2$ symmetry. We first convert every $\boldsymbol{x}_a = (\beta, \phi)$ to a unit vector $\boldsymbol{x}_a = (x, y, z)$. To flip only the first axis, the irrep is $\rho_{1,0,0} = \rho_1^x \boxtimes \rho_0^y \boxtimes \rho_0^z$, which assigns a $\mathbb{R}^{1\times 1}$ matrix to each g. The input data has three channels, x, y, and z. We define the representation to be $\rho_{\texttt{in}} = \rho_{1,0,0} \oplus \rho_{0,1,0} \oplus \rho_{0,0,1}$, a diagonal matrix $\rho_{\texttt{in}}(g) \in \mathbb{R}^{3\times 3}$ as shown in Figure 1 such that each C_2 acts only on a single channel.

3.2. Equivariant Model

The G-equivariant MLPs have input $\boldsymbol{x}_a \in \mathbb{R}^3$ (acted on by $\rho_{\texttt{in}}$ as described in Section 3.1) and 2 hidden layers. These hidden layers are represented with $\rho_{\texttt{hidden}}$ as the *regular representation* of G, pointwise ReLU activation, and 256 hidden units. Then group pooling is applied to achieve invariance, where $\rho_{\texttt{out}}$ is the *trivial representation*. This latent vector is concatenated with the $\boldsymbol{x}_p$ features to feed into an unconstrained network also with 2 hidden layers, ReLU activation, and 256 hidden units. We use a learning rate of $1e-5$ throughout. The output is a drag coefficient $y \in \mathbb{R}$.

4. Experiments

We evaluate the sample efficiency and performance of our invariant networks against unconstrained networks, which have the same architecture as in Section 3.2 except with $\rho_{\texttt{in}}$ as the trivial representation. Training data is generated with Sheridan et al.'s improved version of the RSM toolkit. We define the same upper and lower bounds for the feature vector used in Paul et al. (2023), which are selected for each RSM simulation sample using Latin Hypercube sampling (LHS) (McKay et al., 2000) as detailed in Appendix A, Table 2.

4.1. Cube Satellite

We generate 51,000 samples with atomic hydrogen particles. We compare 5 different symmetry constraints listed in Table 1 over varying sizes of a training set, though the task should

Table 1: Test RMSEs after training for 20 epochs: symmetry constraints on network vs # of training samples (100, 500, etc). C_2^x is cyclic group reflecting the x axis, C_4^{xz} is rotations on the xz plane. Size of the test set was kept at 500 samples for cube, 100 for GRACE. Results averaged over 2 runs. N/A for symmetries that mismatch satellite geometry.

	Cube, # training samples				GRACE, # training samples	
Symmetry	**100**	**500**	**10000**	**50000**	**100**	**500**
Unconstrained	7.63 ± 0.01	0.659 ± 0.0	0.390 ± 0.0	0.267 ± 0.0	6.705 ± 0.032	3.777 ± 0.136
C_2^y	7.23 ± 0.2	1.113 ± 0.0	0.368 ± 0.0	0.274 ± 0.0	6.391 ± 0.003	3.224 ± 0.0
$C_2^x \times C_2^y$	7.02 ± 0.04	1.032 ± 0.0	0.365 ± 0.0	0.274 ± 0.0	N/A	N/A
$C_2^x \times C_2^y \times C_2^z$	6.99 ± 0.03	0.902 ± 0.0	0.368 ± 0.0	0.277 ± 0.0	N/A	N/A
C_4^{xz}	8.88 ± 0.001	0.881 ± 0.0	0.438 ± 0.0	0.321 ± 0.0	N/A	N/A
O_h^{xyz}	7.865 ± 0.016	0.584 ± 0.0	0.458 ± 0.0	0.321 ± 0.0	N/A	N/A

theoretically have octahedral O_h symmetry. Results show that 1) weaker constraints like $C_2 \times C_2 \times C_2$ actually have better performance with extremely limited data, 2) O_h outperforms other constraints with a slightly larger data set of 500 samples, and 3) weaker constraints outperform O_h with very large amounts of data. We believe this is because O_h constraints, shown in Figure 2 (bottom), may not actually match the ground truth function shown in Figure 2 (top), but may have advantages for rougher approximations with less data. This mismatch could be either caused by simulations not capturing the real world symmetry, or incorrect assumptions about the domain's properties.

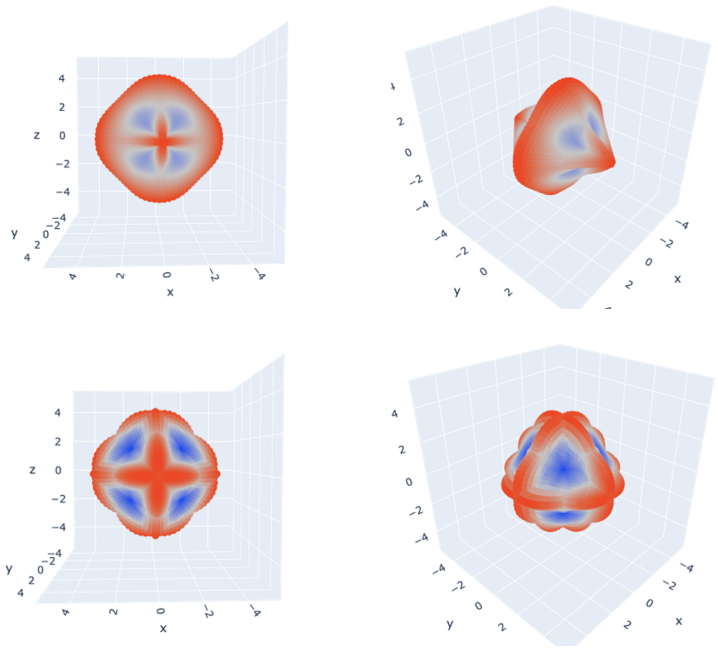

Figure 2: (Top left and top right): Ground truth simulated drag coefficient of a cube with fixed arbitrary $\boldsymbol{x}_p$ variables, where the angle from the origin represents $\boldsymbol{x}_a$, with each point scaled and colored by the magnitude of the drag coefficient at that attitude. It has $C_2 \times C_2 \times C_2$ invariance, reflections across each principal axis. (Bottom left and bottom right): Learned approximation of the cube's drag coefficient, constrained by O_h symmetry.

4.2. GRACE Satellite

Because of computational costs for the more complex geometry of GRACE (pictured in Appendix A, Figure 3), we only generated 576 samples with atomic hydrogen particles. With 64 Intel E5-2680 v4 CPUs each with 32GB of RAM, this amount of data took 17 hr 40 min to generate, showing a need for sample efficient methods. GRACE only has C_2 invariance across the y axis. Results in Table 1 show the benefit of enforcing this constraint, as it reached lower test RMSE for both small amounts and slightly larger amounts of data.

5. Conclusion

We demonstrate the advantages of constraining drag coefficient models to the symmetry of the satellite, showing its benefits when dealing with little training time and data. This is especially useful when learning from TPMC simulations which are computationally expensive and infeasible for satellite orbit projections themselves.

Future work should examine more extensive hyperparameter searches and comparisons, along with longer training times with similar constraints. More satellites with symmetry should be examined, and we encourage designing satellites with symmetry in mind for better in-flight drag modeling. We also see potential to utilize intrinsic equivariance, where the constrained neural network could overcome small imperfections (such as the bumps for sensors present on either side of GRACE) to model the underlying satellite symmetry (Wang et al., 2023). Additionally, analysis should be performed on why the cube drag coefficient's symmetry was different than the octahedral symmetry of the cube, which could inform future work about geometric modeling of particle simulations like TPMC more generally.

Acknowledgments

We sincerely thank Dr. Smriti Nandan Paul and Logan Sheridan for providing resources and documentation to use the RSM toolkit. This work is supported in part by NSF 2107256, NSF 2134178, and Northeastern University Undergraduate Research and Fellowships' PEAK Base Camp award.

References

Carlo Cercignani and Maria Lampis. Kinetic models for gas-surface interactions. *Transport theory and statistical physics*, 1(2):101–114, 1971.

Gabriele Cesa, Leon Lang, and Maurice Weiler. A program to build E(N)-equivariant steerable CNNs. In *International Conference on Learning Representations (ICLR)*, 2022.

Paul A Chambre and Samuel A Schaaf. *Flow of rarefied gases*, volume 4971. Princeton University Press, 1961.

Taco Cohen and Max Welling. Group equivariant convolutional networks. In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 2990–2999, New York, New York, USA, 20–22 Jun 2016. PMLR.

Marc Finzi, Max Welling, and Andrew Gordon Wilson. A practical method for constructing equivariant multilayer perceptrons for arbitrary matrix groups. In *International Conference on Machine Learning*, 2021.

Donald J Kessler, Nicholas L Johnson, JC Liou, and Mark Matney. The kessler syndrome: implications to future space operations. *Advances in the Astronautical Sciences*, 137(8): 2010, 2010.

Michael D McKay, Richard J Beckman, and William J Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1):55–61, 2000.

Piyush M. Mehta, Andrew Walker, Earl Lawrence, Richard Linares, David Higdon, and Josef Koller. Modeling satellite drag coefficients with response surfaces. *Advances in Space Research*, 54(8):1590–1607, 2014a. ISSN 0273-1177. doi: https://doi.org/10.1016/j.asr.2014.06.033.

Piyush M Mehta, Andrew Walker, Craig A McLaughlin, and Josef Koller. Comparing physical drag coefficients computed using different gas–surface interaction models. *Journal of Spacecraft and Rockets*, 51(3):873–883, 2014b.

Smriti Nandan Paul, Phillip Logan Sheridan, Richard J. Licata, and Piyush M. Mehta. Stochastic modeling of physical drag coefficient – its impact on orbit prediction and space traffic management. *Advances in Space Research*, 72(4):922–939, 2023. ISSN 0273-1177. doi: https://doi.org/10.1016/j.asr.2023.06.006.

Phillip Logan Sheridan, Smriti Nandan Paul, Guillermo Avendaño-Franco, and Piyush M. Mehta. Updates and improvements to the satellite drag coefficient response surface modeling toolkit. *Advances in Space Research*, 69(10):3828–3846, 2022. ISSN 0273-1177. doi: https://doi.org/10.1016/j.asr.2022.02.044.

Byron D Tapley, S Bettadpur, Mo Watkins, and Ch Reigber. The gravity recovery and climate experiment: Mission overview and early results. *Geophysical research letters*, 31 (9), 2004.

Andrew Walker and Earl Lawrence. Response surface modeling tool suite, version 1.x, version 00, 7 2016.

Dian Wang, Jung Yeon Park, Neel Sortur, Lawson LS Wong, Robin Walters, and Robert Platt. The surprising effectiveness of equivariant models in domains with latent symmetry. In *The Eleventh International Conference on Learning Representations (ICLR)*, 2023.

Rui Wang, Robin Walters, and Rose Yu. Incorporating symmetry into deep dynamics models for improved generalization. In *The Ninth International Conference on Learning Representations (ICLR)*, 2021.

Appendix A. Additional Diagrams

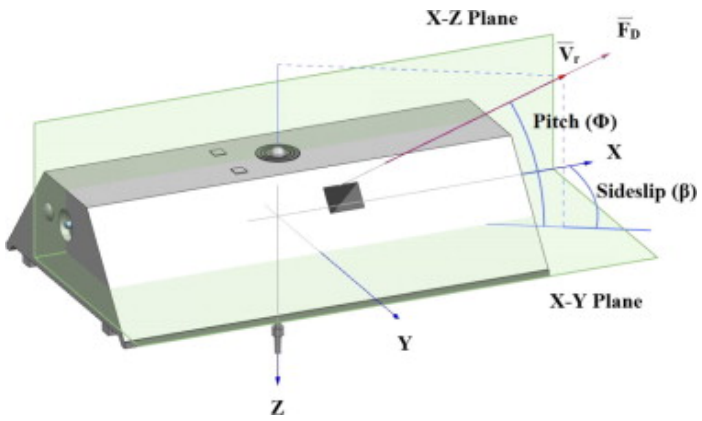

Figure 3: 3D model of the GRACE satellite (Mehta et al., 2014a). There is C_2 symmetry represented as reflections across the XZ plane.

Table 2: Upper and lower bounds for LHS sampling

Independent Variable	**Lower Bound**	**Upper Bound**
x_1	7250.0m/s	8000.0m/s
x_2	100.0K	2000.0K
x_3	200.0K	2000.0K
x_4	0.0	1.0
x_5	0.0	1.0
β	0°	360°
ϕ	-90°	90°