



Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 72 (2023) 922–939

ADVANCES IN  
SPACE  
RESEARCH  
(a COSPAR publication)  
[www.elsevier.com/locate/asr](http://www.elsevier.com/locate/asr)

# Stochastic modeling of physical drag coefficient – Its impact on orbit prediction and space traffic management

Smriti Nandan Paul <sup>a,\*</sup>, Phillip Logan Sheridan <sup>a</sup>, Richard J. Licata <sup>a</sup>, Piyush M. Mehta <sup>a</sup>

<sup>a</sup> Department of Mechanical and Aerospace Engineering, West Virginia University, Morgantown 26506, United States

Received 15 October 2022; received in revised form 23 March 2023; accepted 5 June 2023

Available online 10 June 2023

## Abstract

Ambitious satellite constellation projects by commercial entities and the ease of access to space in recent times have led to a dramatic proliferation of low-Earth space traffic. It jeopardizes space safety and long-term sustainability, necessitating better space domain awareness (SDA). Correct modeling of uncertainties in force models and orbital states, among other things, is an essential part of SDA. For objects in the low-Earth orbit (LEO) region, the uncertainty in the orbital dynamics mainly emanate from limited knowledge of the atmospheric drag-related parameters and variables. In this paper, which extends the work by Paul et al. (2021), we develop a feed-forward deep neural network model for the prediction of the satellite drag coefficient for the full range of satellite attitude (i.e., satellite pitch  $\in (-90^\circ, +90^\circ)$  and satellite yaw  $\in (0^\circ, +360^\circ)$ ). The model simultaneously predicts the mean and the standard deviation and is well-calibrated. We use numerically simulated physical drag coefficient data for training our neural network. The numerical simulations are carried out using the test particle Monte Carlo method using the *diffuse reflection with incomplete accommodation* gas-surface interaction model. Modeling is carried out for the well-known CHAllenging Minisatellite Payload (CHAMP) satellite. Finally, we use the Monte Carlo approach to propagate CHAMP over a three-day period under various modeling scenarios to investigate the distribution of radial, along-track, and cross-track orbital errors caused by drag coefficient uncertainty. The key takeaways of this paper are - (a) a constant drag coefficient cannot be used for reliable SDA purposes, and (b) stochastic machine learning models allow for the computation of drag coefficients in a timely manner while providing reliable uncertainty estimates.

© 2023 COSPAR. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Satellite drag coefficient; Orbit uncertainty quantification; Neural network

## 1. Introduction

Since the launch of the first artificial satellite, Sputnik, and up until recently, the near-Earth space environment saw a nearly balanced, steady growth of objects. However, recent advancements in commercially viable space technologies, satellite mega-constellation launches from private aerospace companies, and access to launch vehicles offering secondary payload services have led to unsustainable pop-

ulation growth. Lemmens and Letizia (2020) investigate conjunction events for the European Space Agency (ESA) satellites at low altitude LEO. As highlighted in their study, an increasing and significant portion of those close encounters are due to satellite constellations and small satellites. The LEO population growth, likely to go unabated, warrants better modeling of dynamical uncertainties and a more accurate prediction of orbital errors to make more informed decisions about space situational awareness (SSA) functions such as satellite conjunction occurrence and maneuvers for collision avoidance. The conservative perturbation forces are well-modeled, and the prime source of dynamical uncertainty for an LEO object is the

\* Corresponding author.

E-mail addresses: [smritinandan.paul@mail.wvu.edu](mailto:smritinandan.paul@mail.wvu.edu) (S.N. Paul), [rjlicata@mix.wvu.edu](mailto:rjlicata@mix.wvu.edu) (R.J. Licata), [piyush.mehta@mail.wvu.edu](mailto:piyush.mehta@mail.wvu.edu) (P.M. Mehta).

atmospheric drag. Other non-conservative forces, such as solar radiation pressure (SRP), can also be a source of significant dynamical uncertainty for high area-to-mass ratio (HAMR) objects, but this study focuses only on “typical” space objects which are low area-to-mass ratio (LAMR).

For a satellite with mass  $m$ , the acceleration due to atmospheric drag is given by the following commonly accepted equation:

$$\vec{a}_D = -\frac{1}{2} \rho \frac{C_D A_{proj}}{m} v_{rel} \vec{v}_{rel} \quad (1)$$

where  $\rho$  is the atmospheric density,  $C_D$  is the drag coefficient,  $A_{proj}$  is the projected area of the satellite perpendicular to the flow direction,  $\vec{v}_{rel}$  is the velocity of the satellite relative to the atmosphere, and  $v_{rel}$  is the magnitude of  $\vec{v}_{rel}$ . All parameters listed on the right-hand side in Eq. 1 can have uncertainties. The uncertainty in mass results from the unknown characteristics of space debris/classified objects or through the act of orbital/attitude maneuvers for known active space objects. The uncertainty in the relative velocity  $\vec{v}_{rel}$  arises from the uncertainty in local thermospheric winds, which can be as high as several hundreds of meters per second (March et al., 2019; Drob et al., 2008). Barring a perfectly spherical satellite, the uncertainty in the projected area emanates from missing/uncertain satellite attitude data and uncertainty in the computation of relative velocity vector direction. Much of the literature focuses on modeling the atmospheric density  $\rho$ , which has a complex dependency upon parameters such as geomagnetic indices, solar flux, the composition of the atmosphere, data measured by onboard satellite instrumentation, and others. These parameters themselves have uncertainties, which ultimately equate to uncertainty in the atmospheric density estimation. Instead of focusing on the prevalent density modeling, the current paper focuses on modeling the drag coefficient, a parameter that captures the interaction between the satellite surface and the atmospheric particles. Like density, modeling the drag coefficient is an involved task because of its dependency on atmospheric composition, satellite and atmospheric temperatures, and atmospheric winds.

Broadly speaking, drag coefficient modeling falls under one of the following three categories - fixed, fitted, and physical (Mehta et al., 2022). When using the fixed approach, the drag coefficient is considered to be constant. In the fitted approach, the drag coefficient is estimated using a filtering method as part of an orbit determination process. In this study, our focus is on modeling the physical drag coefficient, which is determined by simulating the exchange of energy and momentum between the surface of the spacecraft and free-stream atmospheric particles (Chambre and Schaaf, 1961). The most common practice for physical drag coefficient determination is to use computationally expensive numerical methods such as the Panel method, Ray-tracing Panel (RTP) method, Test Particle Monte Carlo (TPMC) method, or the Direct Simulation Monte Carlo (DSMC) method (Mostaza Prieto et al.,

2014). To avoid the high computational cost of the numerical methods, Mehta et al. (2014a) build surrogate models based on Gaussian Process Regression (GPR) (Rasmussen, 2004) for predicting satellite drag coefficient. The authors demonstrate that GPR is able to accurately represent the drag coefficient model on which they are based with root mean square percentage errors below 1% for a number of simple and complex geometries. However, Mehta et al. (2014a) do not carry out uncertainty quantification. Building upon the work by Mehta et al. (2014a), Paul et al. (2021) develop GPR and Monte Carlo Dropout (Gal and Ghahramani, 2016) based feed-forward deep neural network (FFDNN) models for stochastic/probabilistic prediction of the satellite drag coefficient. The authors demonstrate that both models are able to produce reasonably accurate and well-calibrated drag coefficient estimates. However, their models are valid for only a limited attitude range with satellite pitch and yaw varying between  $-10^\circ$  and  $+10^\circ$ . The current paper aims to develop stochastic models for drag coefficient prediction for the full attitude range, i.e., satellite pitch  $\in (-90^\circ, +90^\circ)$  and satellite yaw  $\in (0^\circ, +360^\circ)$ . Additionally, Paul et al. (2021) use a mere 1000 ensemble points for training their machine learning models, which is not sufficient for the full attitude range. Tens of thousands of ensemble training points (Sun et al., 2019a) are desirable to accurately capture the full-attitude drag coefficient variations. Because GPR scales poorly with the data size, it is computationally infeasible to use GPR for the full-attitude drag coefficient modeling. Partially scalable variants of GPR are discussed later in the paper. Machine learning models such as the Monte Carlo Dropout-based FFDNN (Paul et al., 2021) or other variants of the neural network are scalable and can also provide an uncertainty estimate, making them ideal for large-scale full-attitude drag coefficient modeling.

In this paper, we use an in-house developed FFDNN model that directly predicts the mean and the standard deviation. We prefer the direct prediction of the uncertainty against an ensemble approach like the Monte Carlo dropout technique for uncertainty quantification because of the lesser computational costs in the direct prediction method. For any regression, the quality and quantity of the training data are critical to the correctness of the developed machine learning model. For machine learning training purposes in the current study, high-quality drag coefficient data are generated using the numerical TPMC method (Davis, 1960). The gas-surface interaction (GSI) in the TPMC simulations is modeled using the Diffuse Reflection With Incomplete Accommodation (DRIA) model (Walker et al., 2014a; Walker et al., 2014b).

The current paper has a number of desirable goals: (1) the developed drag coefficient prediction model must be quick to evaluate and valid for any satellite orientation, (2) the developed model should be accurate not only in terms of the mean predictions but also provide meaningful and reliable uncertainty estimates, (3) the data size or the number of samples used to train the models should be

sufficiently large to capture the drag coefficient variations with sufficient accuracy, and (4) the developed models should be inductive (as opposed to transductive) so that they can be saved and later re-used in an orbit propagation framework and perform orbital perturbation studies in a computationally efficient manner.

We organize the remainder of this paper into the following sections: Section 2 provides brief background knowledge about the TPMC method and the DRIA GSI model. Section 3 provides details about the data used to train the regression models. We discuss the machine learning model used for the drag coefficient prediction in Section 4. Model calibration, which is a concept closely related to the reliability of the predicted uncertainty, is discussed in Section 5. Prediction performance for different data sizes is analyzed in Section 6. Section 7 highlights difficulties in using scalable Gaussian processes for large-scale drag coefficient modeling. We investigate the effects of drag coefficient uncertainty on the orbital state uncertainties in Section 8. Section 9 discusses the proposed approach's applicability and limitations. Finally, we summarize the paper and provide the conclusions in the last section.

## 2. Background - TPMC and the DRIA GSI model

The TPMC is a numerical method for computing the physical drag coefficient in the free molecular flow (FMF) regime. In TPMC, test particles representing actual molecules are sequentially fired into the computational domain. Each particle is fired with a probabilistically determined velocity. The TPMC assumes that molecules do not collide with one another, which speeds up computations while maintaining accuracy on par with other Monte Carlo techniques like the DSMC. The TPMC method is versatile because it can simulate different GSI models and easily handle flows with complex boundaries.

As stated, our TPMC simulations are carried out using the DRIA GSI model. In the DRIA GSI model, the reflected particles have a diffuse angular distribution, based on Knudsen's cosine law (Knudsen, 1916). The particles may exchange energy with the surface depending on the value of the energy accommodation coefficient. For more details on the DRIA GSI model, refer to Walker et al. (2014a), Walker et al. (2014b), Pilinski et al. (2010), Moe et al. (2004), Sutton (2009).

## 3. Input data for the machine learning models

The training data for our predictive models are generated by numerical computation of the physical drag coefficient using the TPMC method, which is implemented using the West Virginia University (WVU) Response Surface Modeling (RSM) toolkit (Sheridan et al., 2022). The WVU RSM toolkit is an open-source software package and can be accessed at [https://github.com/ASSISTLaboratory/WVU\\_RSM\\_Suite](https://github.com/ASSISTLaboratory/WVU_RSM_Suite). In this paper, drag coefficient modeling is carried out for the following primary LEO

atmospheric constituents - atomic hydrogen ( $H$ ), helium ( $He$ ), atomic nitrogen ( $N$ ), molecular nitrogen ( $N_2$ ), atomic oxygen ( $O$ ), and molecular oxygen ( $O_2$ ). The total drag coefficient can be computed from the drag coefficients of the constituent species.

For the TPMC method with the DRIA GSI model, six independent variables determine the value of the dependent drag coefficient - (i) relative velocity of the satellite,  $v_\infty$ , (ii) satellite surface temperature,  $T_w$ , (iii) local free-stream temperature,  $T_\infty$ , (iv) energy accommodation coefficient,  $\alpha$ , (v) satellite yaw,  $\beta$ , and, (vi) satellite pitch,  $\Phi$ . In this study, the input configurations for the numerical simulations are carefully selected via the Latin Hypercube sampling (LHS) method (McKay et al., 1979). We use a high-fidelity geometry model corresponding to the CHAllenging Minisatellite Payload (CHAMP) satellite for our analysis. A total of 50,000 LHS design points are selected for each of the species  $H, He, N, N_2, O, O_2$  for training purposes. The upper and lower bounds defining the LHS design points are given in Table 1.

The number of training samples, i.e., 50,000, is determined through a data size sensitivity analysis, which is discussed later. Besides the training data points, a different set of 50,000 LHS points, using the same bounds as that of Table 1, are constructed for validation/testing purposes for each of the six species.

## 4. Constrained dual prediction of mean and standard deviation using FFNN With NLPD loss function

One of the most popular supervised machine learning techniques roughly based on the working of a human brain is the feed-forward neural network (FFNN). It is characterized by an input layer, an output layer, and layers in-between called hidden layers, where each layer is composed of *neurons* or *nodes*. An FFNN, especially one with multiple hidden layers (the so-called feed-forward deep neural network or FFDNN), often consists of a large number of parameters (*weights* and *biases*) that control the function mapping from one layer to the next. The depiction of the mapping for a dummy 3-layer FFNN is shown in Fig. 1. The parameters  $a_i^{(j)}$  in Fig. 1 are called *activation units*, and their expressions are given in Eqs. (2)–(5). The functions  $g_1$  and  $g_2$  appearing in Eqs. (2)–(5) are some user-defined *activation functions*, and  $\tilde{\theta}^{(j)}$  denotes the matrix of weights controlling function mapping from layer  $j$  to layer

Table 1  
Upper and lower bounds defining the LHS design points.

Independent Variables	Lower Bound	Upper Bound
$v_\infty$	7250.0 m/s	8000.0 m/s
$T_w$	100.0 K	2000.0 K
$T_\infty$	200.0 K	2000.0 K
$\alpha$	0.0	1.0
$\beta$	0°	360°
$\Phi$	-90°	90°

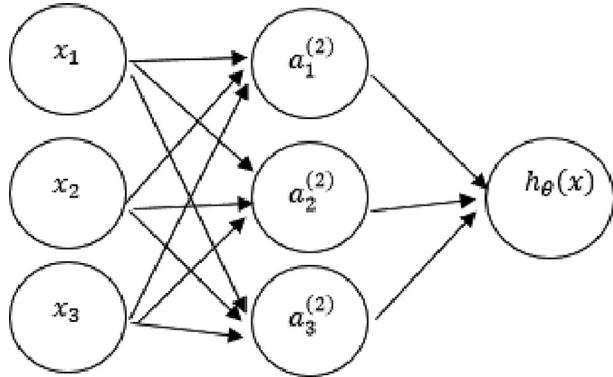


Fig. 1. Neural network mapping for a dummy 3-layer case with a three dimensional input and a univariate output.

$(j+1)$ . The elements of matrix  $\tilde{\theta}^{(j)}$  are determined from the optimization of a user-defined  $\tilde{\theta}^{(j)}$ -dependent cost function. For more details on FFNN, see Svozil et al. (1997).

$$a_1^{(2)} = g_1(\tilde{\theta}_{10}^{(1)} + \tilde{\theta}_{11}^{(1)}x_1 + \tilde{\theta}_{12}^{(1)}x_2 + \tilde{\theta}_{13}^{(1)}x_3) \quad (2)$$

$$a_2^{(2)} = g_1(\tilde{\theta}_{20}^{(1)} + \tilde{\theta}_{21}^{(1)}x_1 + \tilde{\theta}_{22}^{(1)}x_2 + \tilde{\theta}_{23}^{(1)}x_3) \quad (3)$$

$$a_3^{(2)} = g_1(\tilde{\theta}_{30}^{(1)} + \tilde{\theta}_{31}^{(1)}x_1 + \tilde{\theta}_{32}^{(1)}x_2 + \tilde{\theta}_{33}^{(1)}x_3) \quad (4)$$

$$a_1^{(3)} = h_\theta(x) = g_2(\tilde{\theta}_{10}^{(2)} + \tilde{\theta}_{11}^{(2)}a_1^{(2)} + \tilde{\theta}_{12}^{(2)}a_2^{(2)} + \tilde{\theta}_{13}^{(2)}a_3^{(2)}) \quad (5)$$

In this paper, we use a variant of the FFNN that simultaneously predicts the mean and standard deviation of the drag coefficient as a function of the input features (see Licata and Mehta (2022)). For networks modeling the drag coefficient associated with the DRIA GSI model, we have the following eight features:  $v_\infty, T_w, T_\infty, \alpha, \sin \beta, \cos \beta, \sin \Phi, \cos \Phi$ . Instead of a univariate output, we predict an output with shape [2, 1]. The first output node represents the mean drag coefficient and the second output node represents the corresponding standard deviation. For the proposed variant of the FFNN, the input training data has an array structure of [number of samples, feature dimension], where the feature dimension is eight for DRIA. The output training data is augmented with zeros to have an array structure of [number of samples, 2, 1]. Since standard deviation values can only be non-negative, we constrain the second output node to be non-negative using the *softplus function* so that meaningful uncertainties are predicted. One may also use alternatives such as the *absolute value function* to impose the positivity constraint on the standard deviation prediction, but it may lead to unstable predictions because of the non-smooth first derivative at zero. For training purposes, we use the *negative logarithm of the probability density* (NLPD) loss function, given as:

$$\text{loss}_{\text{NLPD}} = \frac{1}{n_t} \sum_{i=1}^{n_t} \left[ \log \hat{\sigma}_i^2 + \frac{\|\mathbf{y}_i - \hat{\mu}_i\|^2}{\hat{\sigma}_i^2} + \log 2\pi \right] \quad (6)$$

where  $(\mathbf{x}_i, \mathbf{y}_i)_{i=1,\dots,n_t}$  represent the training data set,  $\hat{\mu}_i$  represents the prediction mean, and  $\hat{\sigma}_i^2$  represents the prediction variance.

Although the training data set only contains a cosmetic set of zeros as standard deviations, we will see later in this paper that the functional form of NLPD is powerful enough to allow the network to learn well-calibrated uncertainty estimates while also producing reasonably accurate mean estimates. The implementation of the model is carried out using the Keras deep learning API (Chollet, 2015), where we define a custom class for modifying the in-built dense layers to meet our constraints. A dummy FFNN with three inputs for dual prediction of the mean and standard deviation of a variable  $y$  is shown in Fig. 2.

## 5. Model calibration

Calibration is the requirement in stochastic modeling that the predicted probabilities approximate the probability of actual events (Camporeale and Carè, 2020). A well-calibrated model (assuming Gaussian distribution), for example, should have around 68% of observations within one standard deviation of the predicted mean, 95% of observations within two standard deviations of the predicted mean, and 99.7% of observations within three standard deviations of the predicted mean. Uncalibrated models tend to be overconfident or underconfident in their predictions, and one should not trust their inferences. A convenient way to check how well a model is calibrated is by looking at its “consistency curve” or “calibration curve”.

Let the expected confidence interval levels be:  $C = [5\%, 10\%, 15\%, \dots, 95\%]$ . The corresponding coefficients defining the uncertainty bounds are then given as:  $\zeta[k] = \sqrt{2} \operatorname{erf}^{-1}(C[k]/100)$ , where erf is the well-known error function. Let  $(\mathbf{x}_{oj}, \mathbf{y}_{oj})_{j=1,\dots,m}$  be the observation data set and let the corresponding predictions be  $(\hat{\mu}_j, \hat{\sigma}_j)_{j=1,\dots,m}$ , where  $\hat{\mu}_j$  represents the mean and  $\hat{\sigma}_j$  represents the standard deviation. Then, the percentage of the observed data set within the lower and upper uncertainty bounds associated with  $C[k]$  is obtained as (Anderson et al., 2020):

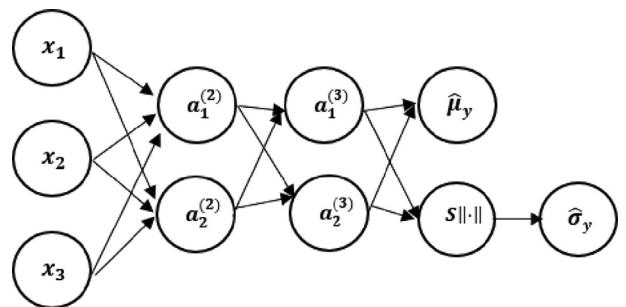


Fig. 2. Dummy feed-forward neural network for dual prediction of the mean and standard deviation. The non-negative constraint on the standard deviation prediction is indicated by the softplus function  $S||\cdot||$ .

$$P[k] = \left[ \sum_{j=1}^m \frac{\mathcal{I}((\mu_j - \zeta[k]\sigma_j) < y_{oj} < (\mu_j + \zeta[k]\sigma_j))}{m} \right] \times 100 \quad (7)$$

where  $\mathcal{I}$  is the indicator function.

The consistency curve mentioned earlier is the plot of  $P$  versus  $C$ . The proximity of the consistency curve to the  $y = x$  line (i.e., a straight line with a slope of  $45^\circ$  and passing through the origin) is used to measure calibration in this study. The consistency curve will perfectly overlap the  $y = x$  line in a perfectly calibrated system (this does not usually happen in real-life models).

In addition to the qualitative performance of the estimated uncertainties, we also investigate the quantitative performance using the mean absolute calibration error (MACE), defined as:

$$MACE = \frac{1}{n_C} \sum_{k=1}^{n_C} |C[k] - P[k]| \quad (8)$$

where  $n_C$  is the number of confidence interval levels.

In this work, in order to improve the model calibration performance, we scale the predicted standard deviation values using the following scaling factor (Laves et al., 2021):

$$s = \sqrt{\frac{1}{n_v} \sum_{i=1}^{n_v} \left[ \frac{\|\mathbf{y}_i - \hat{\mu}_i\|^2}{\hat{\sigma}_i^2} \right]} \quad (9)$$

where  $n_v$  is the number of validation data samples,  $\mathbf{y}_i$  is the TPMC drag coefficient,  $\hat{\mu}_i$  is the predicted mean drag coefficient, and  $\hat{\sigma}_i^2$  is the predicted standard deviation for the  $i^{\text{th}}$  validation sample.

## 6. Prediction performance for different data sizes

In this section, we first simulate drag coefficients for a sphere, assuming that the drag coefficient is attitude-independent. Following that, we carry out drag coefficient simulations for the CHAMP satellite. Because of its complex geometry and drag coefficient's dependence on attitude, we demonstrate that a satellite like the CHAMP necessitates a much larger number of data points for neural network training to achieve sufficient accuracy and reliability.

We perform neural network simulations for the spherical satellite using 10,000 data points for training/validation

purposes for each of the species  $H, He, N, N_2, O, O_2$ . A critical task in the design of neural network architecture is to decide on the values of the model hyper-parameters. These are the parameters that we cannot learn from the training process. In our work, we restrict ourselves to the hyper-parameter optimization (also known as “tuning”) of the following quantities: (i) number of hidden layers, (ii) number of neurons in each hidden layer, (iii) activation functions, (iv) Monte Carlo dropout rates, (v) network optimizer, and (vi) batch size. The traditional practice uses a heuristic or manual approach to select the hyper-parameters. In our work, however, we use the KerasTuner library (O’Malley et al., 2019) to find our near-optimal hyper-parameters. Within KerasTuner, we have the following optimizer options: (i) random search, (ii) Bayesian optimization, and (iii) hyperband; this work uses Bayesian optimization. Table 2 lists the search space used for the hyper-parameter tuning and Table 3 lists the essential tuning parameters.

We carry out the training using the best architecture resulting from the KerasTuner optimization process. The training data set comprises 8,500 samples obtained through a 15:85 split of 10,000 samples, with 15% data samples used for validation and the remaining used for training. The performance of the trained network is assessed on a test data set of size 10,000. The 10,000 test samples are distinct from the training/validation samples and are generated using TPMC in the same way that the training data sets are. Note that the machine learning input dimension is now reduced to four (from eight for the CHAMP satellite) because the sphere is assumed to be rotationally invariant. Fig. 3 shows the neural network prediction results for the test data set. On the x-axis, we have the observed (numerical) drag coefficients; on the left y-axis, we have the predicted mean drag coefficients; on the right y-axis, we have the predicted  $3\sigma$  uncertainty values. The calibration curves corresponding to the prediction shown in Fig. 3 are not included here

Table 3  
Tuning parameters used in the hyper-parameter optimization.

Tuning Parameter	Value
Maximum number of trials	150
Executions per trial	5
Number of initial points	50
Early stop regularization patience	50
Number of epochs	200

Table 2  
Hyper-parameter search space.

Hyper-parameter	Values
Number of hidden layers	[1, 2, 3, ..., 9, 10]
Number of neurons in each hidden layer	[32, 64, 96, ..., 512, 1024]
Activation function for the neurons of each hidden layer	[relu, tanh, sigmoid, softsign, selu, elu, linear]
Dropout rate for each hidden layer	[.05, .10, .15, ..., .75, .80]
Optimizer for neural network training	[rmsprop, adagrad, adam, nadam]
Batch size	[256, 512, 768, ..., 3840, 4096]

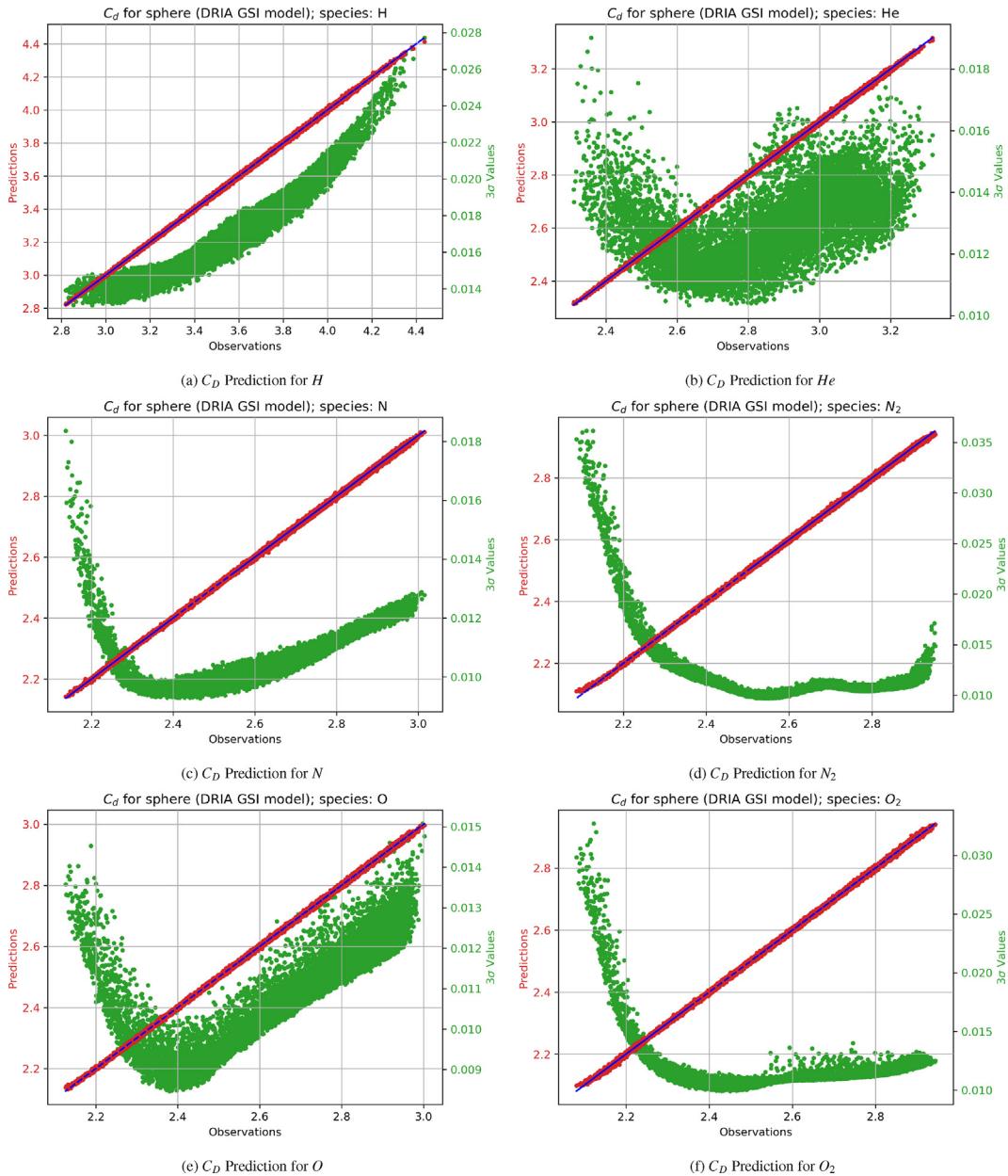


Fig. 3. Comparison of the true and predicted drag coefficients for the test data set for a spherical satellite. In red, we have the predicted means, and in green, we have the uncertainty predictions. Training is carried out using 8,500 samples. The testing is carried out using a different set of 10,000 samples.

for brevity. The RMSE and the MACE values for prediction on the test data set are given in Table 4. For reference, the RMSE results obtained here are comparable to that of the results obtained by Mehta et al. (2014a), where the

authors use Gaussian Process regression and 1000 ensemble points for training purposes. From Fig. 3 and Table 4, we can conclude that 10,000 data points are sufficient for producing accurate and well-calibrated results for a simple geometry like sphere.

Next, we perform simulations for the CHAMP satellite. We perform a sensitivity analysis to determine the appropriate data size for neural network training. We inspect three different data sizes - (1) training with 10,000 samples, (2) training with 20,000 samples, and (3) training with 50,000 samples for each of the species H, He, N, N<sub>2</sub>, O, O<sub>2</sub>. The performance of the trained networks is assessed on test data sets. In addition, the neural network predicted drag coefficients are compared to SPARTA-based drag coefficient data provided by Dr. Christian Siemes of the Delft

Table 4

RMSE and MACE performance on the test data set for all species for a spherical satellite.

Species	RMSE	MACE (%)
H	0.0058	0.2879
He	0.0044	0.4810
N	0.0038	0.9447
N <sub>2</sub>	0.0038	0.5584
O	0.0037	0.7421
O <sub>2</sub>	0.0038	0.8310

University of Technology (March et al., 2021). The accuracy of the neural network predictions compared to the SPARTA-based drag coefficients and the computational costs determine the appropriate data size. For interested readers, Stochastic PArallel Rarefied-gas Time-accurate Analyzer (SPARTA) (Plimpton et al., 2019) is a parallel Direct Simulation Monte Carlo (DSMC) code for performing simulations of low-density gases in 2D or 3D.

We carry out the training for the three data sizes (10,000, 20,000, and 50,000) using the best architecture resulting from the KerasTuner optimization process (based on Tables 2 and 3). Subsequently, we use the trained

networks for drag coefficient prediction for the test data set. The test data set contains 42,500 samples obtained through a 15:85 split of 50,000 samples, with 15% data samples used for validation and the remaining 85% data samples used for testing. The 50,000 validation/test samples are distinct from the training samples. For training using 50,000 samples, the prediction performance on the test data set is shown in Fig. 4. The observed drag coefficients are on the x-axis, the neural network predicted mean drag coefficients are on the left y-axis, and the neural network predicted  $3\sigma$  uncertainty values are on the right y-axis.

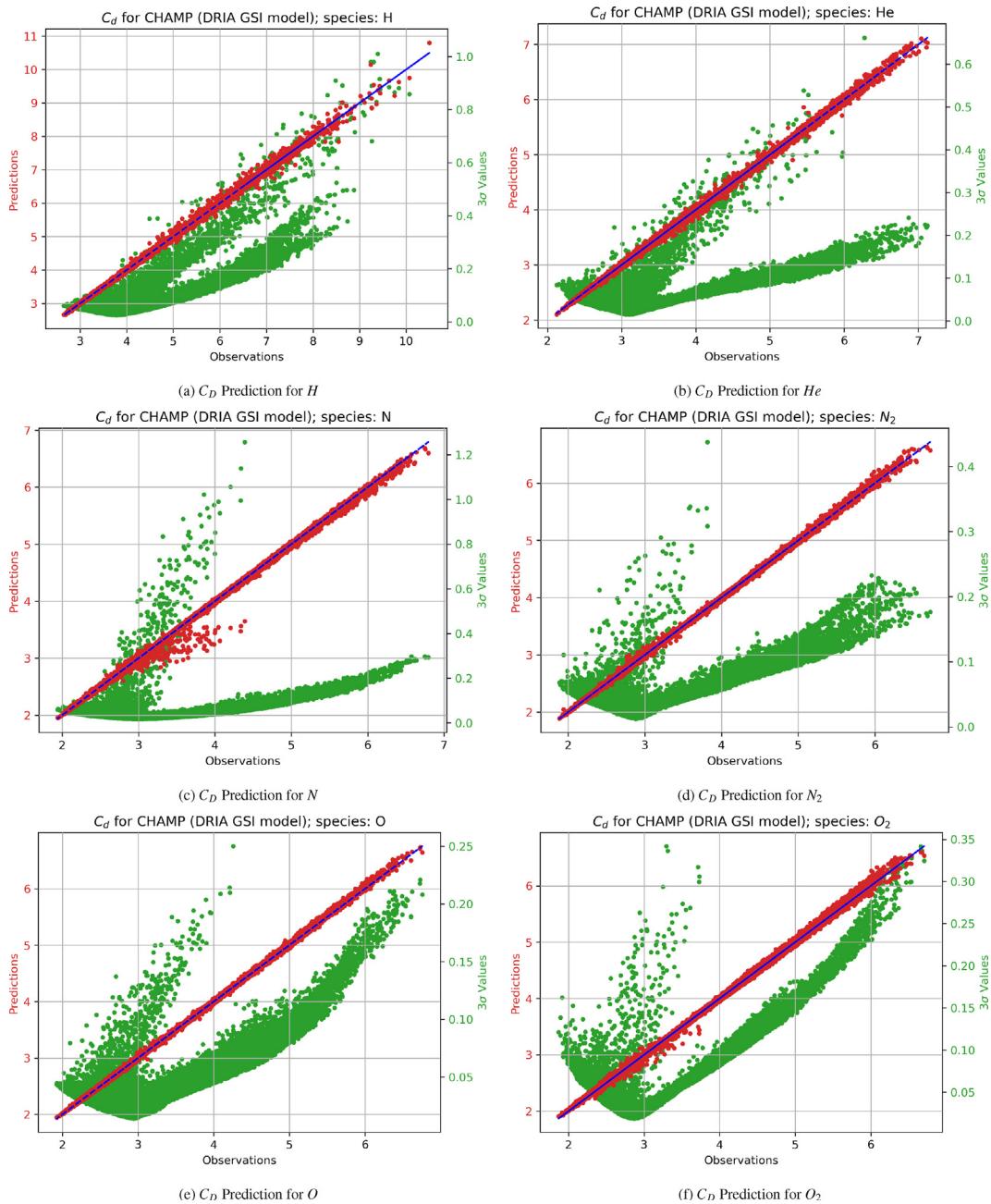


Fig. 4. Comparison of the true and predicted drag coefficients for the test data set for the CHAMP satellite. In red, we have the predicted means, and in green, we have the uncertainty predictions. Training is carried out using 50,000 samples. The testing is carried out using a different set of 42,500 samples.

data set for models trained using 50,000 samples is given in Fig. 5. We have the calibration curve in red, and in blue, we have the reference  $y = x$  line. Also shown in the figure, in green, we have the calibration curve corresponding to the case if we were “not” to scale the neural network predicted standard deviations using Eq. 9. It is evident from Fig. 5 that the neural network models provide well-calibrated estimates after scaling. We do not show prediction performance plots and calibration curves for training using 10,000 and 20,000 samples for brevity. However, for completion, the RMSE and MACE values for all the species for

training using 10,000, 20,000, and 50,000 samples are shown in Table 5.

Next, for comparison of the neural network-predicted drag coefficients with the DSMC-predicted drag coefficients, the drag coefficients of the individual species are combined to compute the total drag coefficient. The total drag coefficient is given as (Mehta et al., 2014b; Walker et al., 2014c):

$$C_D = f_{sc} C_{D_{ads}} + (1 - f_{sc}) C_{D_{surf}} \quad (10)$$

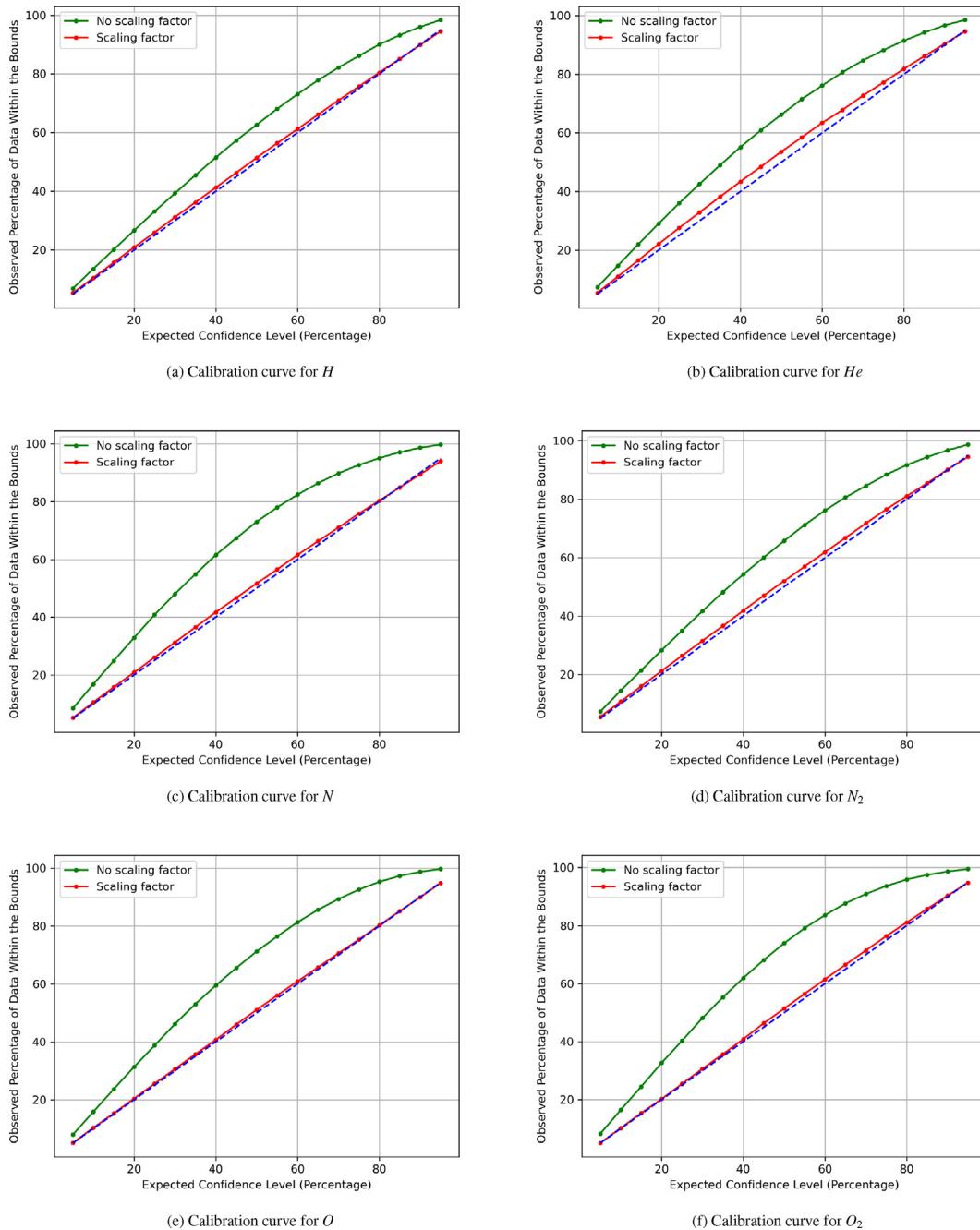


Fig. 5. Calibration curves for the predicted drag coefficient for the test data set for the CHAMP satellite. The prediction models are trained using 50,000 samples.

Table 5

RMSE and MACE performance on the test data set for all species for the CHAMP satellite. Comparison is shown for training carried out using 10 K samples, 20 K samples, and 50 K samples.

Species	RMSE (10 K   20 K   50 K)	MACE (%) (10 K   20 K   50 K)
H	0.0768   0.0686   0.0295	0.9511   0.7778   0.8845
He	0.0337   0.0485   0.0188	3.3381   0.2550   2.2404
N	0.0248   0.0210   0.0250	2.4136   1.5145   1.0279
$N_2$	0.0368   0.0251   0.0159	0.6852   0.4904   1.3170
O	0.0634   0.0475   0.0142	1.9360   1.6784   0.5117
$O_2$	0.0266   0.0445   0.0209	2.2379   1.6260   0.8384

where  $C_{D_{ads}}$  is the total drag coefficient based on a satellite completely covered by the adsorbate (atomic oxygen), and  $C_{D_{surf}}$  is the total drag coefficient based on a clean satellite surface. The weight  $f_{sc}$  is given as (Walker et al., 2014c):

$$f_{sc} = \frac{K_{DRIA} P_o}{1 + K_{DRIA} P_o} \quad (11)$$

where  $K_{DRIA}$  is the Langmuir adsorbate constant for the DRIA model ( $= 1.44 \times 10^6$ ) and  $P_o$  is the partial pressure of atomic oxygen. The adsorbate and the surface drag coefficients are obtained from the drag coefficients of constituent species ( $H, He, N, N_2, O, O_2$ ) using (Walker et al., 2014c):

$$C_{D_{ads/surf}} = \left( \frac{1}{\sum_{k=1}^6 (\chi_k m_k)} \right) \sum_{k=1}^6 (\chi_k m_k C_{D_{ads/surf},k}) \quad (12)$$

where  $\chi_k$  is the mole fraction of species  $k$ ,  $m_k$  is the mass of species  $k$ , and  $C_{D_{ads/surf},k}$  is the drag coefficient for species  $k$ . The adsorbate drag coefficient corresponding to species  $k$ , i.e.,  $C_{D_{ads},k}$ , is obtained by sampling from the distribution predicted by the neural network model for each species with features:  $[v_\infty, 400 \text{ K}, T_\infty, \alpha_{ads}, \sin \beta, \cos \beta, \sin \Phi, \cos \Phi]$ . Similarly, the surface drag coefficient corresponding to species  $k$ , i.e.,  $C_{D_{surf},k}$ , is obtained by sampling from the distribution predicted by the neural network model for each species with features:  $[v_\infty, 400 \text{ K}, T_\infty, \alpha_{surf}, \sin \beta, \cos \beta, \sin \Phi, \cos \Phi]$ . The local free-stream temperature,  $T_\infty$ , is obtained from the NRLMSISE-00 model. We get the satellite relative velocity  $v_\infty$  and the species mole fraction  $\chi_k$  data from DSMC files provided by Dr. Christian Siemes. The energy accommodation coefficients  $\alpha_{ads}$  and  $\alpha_{surf}$  are taken as 0.85 because the DSMC results are generated using an accommodation coefficient of 0.85.

Fig. 6 shows a 24-h comparison between the neural network predicted drag coefficient and the SPARTA-based drag coefficient for four randomly selected days - (a) June 27, 2001, (b) March 27, 2002, (c) January 26, 2004, (d) April 20, 2010 - for the CHAMP satellite. In each of the figures, we also show how the performance varies for models trained using different data sizes. The solid lines show the neural network predicted mean drag coefficient multiplied by the projected area of the CHAMP satellite. Blue color:

the neural networks are trained using 10,000 samples; red color: the neural networks are trained using 20,000 samples; green color: the neural networks are trained using 50,000 samples. In black, we have the DSMC-based drag coefficient multiplied by the projected area of the CHAMP satellite. The shaded green area shows the neural network predicted  $3\sigma$  uncertainty multiplied by the projected area for models trained using 50,000 samples. Uncertainties are not shown for models trained with 10,000 or 20,000 samples for clarity/readability. As evident from Fig. 6, the appropriate number of training samples is 50,000, as using 50,000 training samples results in much better accuracy (determined by closeness to the DSMC results) in comparison to 10,000 or 20,000 training samples.

## 7. Scalable Gaussian process modeling for large-scale drag coefficient modeling - difficulties

Traditional Gaussian Processes provide accurate mean drag coefficient predictions and highly calibrated uncertainty estimates. But they do not scale well with the number of samples or output dimension. Another drawback of the traditional GPR is that the trained models are often too large to be saved, making reusability a challenging task. Despite using powerful computers (GPU-enabled/multiple cores), a GPR model training with 50,000 training samples is unsuccessful due to memory constraints. GPR model training with 10,000/20,000 training samples is feasible, but the saved models are in gigabytes, rendering them useless for later use in orbit uncertainty propagation studies. Therefore, we explore scalable GPR techniques, which are approximate but faster versions of the full-scale Gaussian Processes. Based on the literature review, we explore the following three scalable Gaussian Processes - (1) Parametric Gaussian Processes (PGP) (Raissi et al., 2019), (2) GPflow-based Sparse Variational Gaussian Process (SVGP) (van der Wilk et al., 2020) and (3) local approximate Gaussian Process (laGP) regression (Sun et al., 2019b). For the PGP regression, the training is computationally expensive, lacks sufficient prediction accuracy, and reusability is a challenging task. Therefore, we do not give it further consideration. For the GPflow-based SVGP, reusability and accuracy are the issues. Therefore, we do not recommend using it. The primary challenge with laGP is that the trained approximate models cannot be saved for later use, as it is a transductive model. We seek

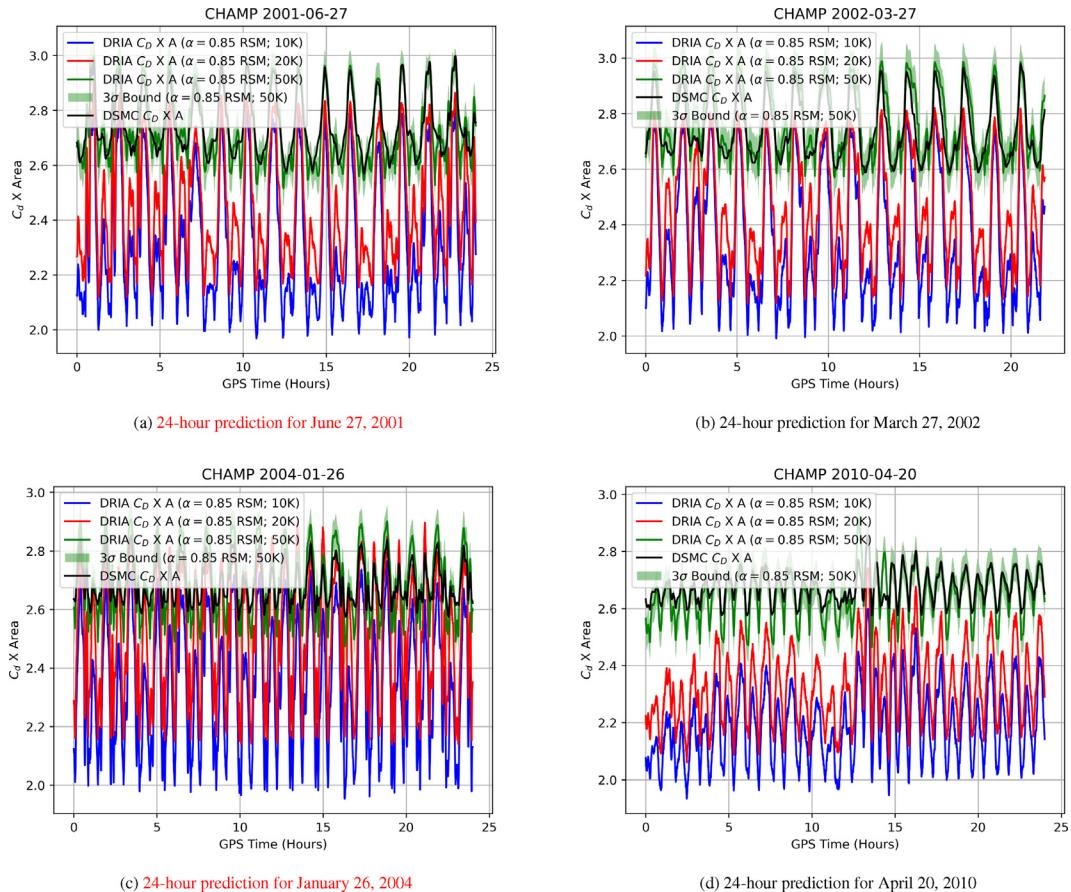


Fig. 6. CHAMP  $C_D \times \text{Area}$  for four select dates for models trained using 10,000, 20,000, and 50,000 samples. Comparison of neural network predicted drag coefficient with SPARTA-based drag coefficient.

an inductive model for reusability purposes. Based on our investigation, we conclude that scalable Gaussian Processes are not ideal for large-scale drag coefficient modeling, where a primary requirement is that the trained models be saved and re-loaded later for orbit uncertainty propagation investigations. An alternative, such as the neural network, is more suitable for our application.

## 8. Effect of physical drag coefficient modeling uncertainties on orbital state uncertainties

In this section, first, we demonstrate the importance of physical drag coefficient modeling for orbit propagation. Following that, we study the effects of drag coefficient uncertainties on orbital state uncertainties.

### 8.1. Demonstration of the importance of physical drag coefficient modeling

In orbit propagation, drag coefficients are often assumed to be constants for simplicity or lack of a computationally efficient method of calculating drag coefficients. To demonstrate the drawback of this assumption and the need for modeling the physical drag coefficient, we consider six test case studies, which are detailed in [Table 6](#). In each

of the six cases, we consider two scenarios: (i) in the first scenario, we propagate an object with a variable drag coefficient which we compute using the machine learning models, and (ii) in the second scenario, the same object is propagated with a constant drag coefficient obtained by averaging the drag coefficient values along the orbit from the first scenario. We then investigate the temporal evolution of the along-track difference between the two scenarios.

The test case objects are assumed to be in a high-inclination, near-zero eccentricity orbit with an altitude of around 400 km. The initial orbital elements for all the test cases are the same and are given in [Table 7](#). Relevant simulation parameters are given in [Table 8](#). It is important to note that the selected propagation period corresponds to a geomagnetic storm during a solar maximum. For orbit integration, we use a modified version of Dormand and Prince's Runge–Kutta Method ([Dormand and Prince, 1980](#)) (also referred to as the 'RK45' integrator in Python's `scipy.integrate` package ([Virtanen et al., 2020](#))). The modified integrator uses a constant integration step size of 10 s rather than striving for specified absolute and relative tolerances. This modification was made because a variable step size integrator takes a long time to converge in the presence of a stochastic drag coefficient, whose value

Table 6

Description of test cases to motivate the importance of physical drag coefficient modeling. The parameter ‘ $D$ ’ is the time in days since the initial epoch of orbit propagation. The arguments for the sine and cosine functions in the definition of pitch and yaw are in radians; however, pitch/yaw themselves are in degrees.

Case Number	Object	Drag Coefficient	Attitude	Cross-Sectional Area
Case I-a	Sphere	Physical $C_{D_i}$ from machine learning	N/A	Constant
Case I-b	Sphere	Constant $C_D = \frac{1}{(n_{steps})} \sum_{orbit} C_{D_i} s$ (case I-a)	N/A	Constant
Case II-a	CHAMP	Physical $C_{D_i}$ from machine learning	Pitch = yaw = $0^\circ$	Constant
Case II-b	CHAMP	Constant $C_D = \frac{1}{(n_{steps})} \sum_{orbit} C_{D_i} s$ (case II-a)	Pitch = yaw = $0^\circ$	Constant
Case III-a	CHAMP	Physical $C_{D_i}$ from machine learning	Pitch = $\sin(100D)^\circ$ , yaw = $5 \cos(100D)^\circ$	Attitude-dependent
Case III-b	CHAMP	Constant $C_D = \frac{1}{(n_{steps})} \sum_{orbit} C_{D_i} s$ (case III-a)	Pitch = $\sin(100D)^\circ$ , yaw = $5 \cos(100D)^\circ$	Attitude-dependent
Case IV-a	CHAMP	Physical $C_{D_i}$ from machine learning	Pitch = $90 \sin(100D)^\circ$ , yaw = $180(1 + \cos(100D))^\circ$	Attitude-dependent
Case IV-b	CHAMP	Constant $C_D = \frac{1}{(n_{steps})} \sum_{orbit} C_{D_i} s$ (case IV-a)	Pitch = $90 \sin(100D)^\circ$ , yaw = $180(1 + \cos(100D))^\circ$	Attitude-dependent
Case V-a	CHAMP	Physical $C_{D_i}$ from machine learning	Pitch = $\sin(100D)^\circ$ , yaw = $5 \cos(100D)^\circ$	Attitude-dependent
Case V-b	CHAMP	Constant $C_D = \frac{1}{(n_{steps})} \sum_{orbit} C_{D_i} s$ (case V-a)	N/A	Constant
Case VI-a	CHAMP	Physical $C_{D_i}$ from machine learning	Pitch = $90 \sin(100D)^\circ$ , yaw = $180(1 + \cos(100D))^\circ$	Attitude-dependent
Case VI-b	CHAMP	Constant $C_D = \frac{1}{(n_{steps})} \sum_{orbit} C_{D_i} s$ (case VI-a)	N/A	Constant

Table 7

Keplerian elements defining the initial position of the satellites.

Orbital Element	Values
Semi-major axis	6778136.3000 m
Eccentricity	$2.2150 \times 10^{-3}$
Inclination	$87.2193^\circ$
True anomaly	$274.4887^\circ$
Argument of perigee	$85.6397^\circ$
Right ascension of ascending node	$206.9785^\circ$

Table 8

Simulation parameters for the test cases.

Parameter	Values/Description
Initial epoch	00:00:00 UT, November 20, 2003
Propagation period	3 days
Perturbations	$J_2$ , atmospheric drag
Atmospheric density model	NRLMSISE-00 (Picone et al., 2002)
Source for $a_p, A_p$ , $F10.7$ solar radio flux	Celestrak (Tapping (2013), Matzka et al. (2021))
Satellite mass	500 kg

changes in every internal adjustment of a single call of the step size. The cross-sectional areas of the objects for test cases I and II are taken as  $0.770981 \text{ m}^2$ . For test cases III, IV, V-a, and VI-a, cross-sectional area (attitude-dependent) is obtained by applying SciPy’s *LinearNDInterpolator* (Virtanen et al., 2020) to an area look-up table. The constant cross-sectional areas for test cases V-b and VI-b are obtained by averaging the cross-sectional area values along the orbit from test cases V-a and VI-a, respectively.

For all six test cases, for the machine learning-based drag coefficient computations, we use the predicted mean drag coefficients (and ignore the predicted standard devia-

tions). For the drag coefficient computation, we use Eq. 10, where the satellite surface temperature ( $T_w$ ) is taken as 400 K, the atmospheric translation temperature ( $T_\infty$ ) is obtained from the NRLMSISE-00 density model, the adsorbate energy accommodation coefficient ( $\alpha_{ads}$ ) is taken as 1 (Mehta et al., 2022), and the surface energy accommodation coefficient ( $\alpha_{surf}$ ) is obtained as (Mehta et al., 2022):

$$\alpha_{surf} = \frac{3\mu}{(1 + \mu)^2}; \quad \mu = \frac{\sum_{k=1}^6 (\chi_k m_k)}{m_{surf}} \quad (13)$$

where  $\chi_k$  is the mole fraction of species  $k$ ,  $m_k$  is the mass of species  $k$ , and  $m_{surf}$  is the mass of a particle that composes the surface lattice (=263.3223 amu (Mehta et al., 2022)). Mole fractions for the species  $H, He, N, N_2, O, O_2$  are obtained from the NRLMSISE-00 density model.

Fig. 7 shows the evolution of along-track errors for the six cases. Test cases I, II, III, IV, V, and VI are shown in red, green, blue, black, magenta, and cyan, respectively. The history of along-track errors shows that assuming a constant drag coefficient can result in errors of a kilometer or more at the end of the three-day propagation period.

## 8.2. Orbital state uncertainties

We use the well-known Monte Carlo approach for orbit uncertainty propagation for the CHAMP satellite. Several different scenarios are investigated, where for each case, we use a total of 500 Monte Carlo runs and investigate the error distribution at the end of three days of orbit propagation. A top-level description of the Monte Carlo approach is given in Table 9. Uncertainty investigations are carried out for different (a) drag coefficient models, (b)

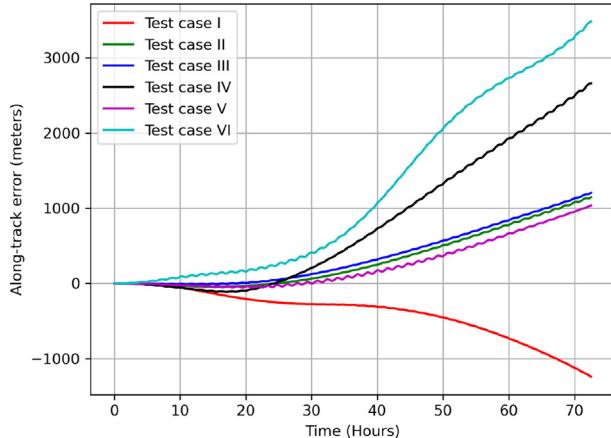


Fig. 7. Evolution of along-track error for different cases, demonstrating the importance of physical drag coefficient modeling.

spatiotemporal correlations for the drag coefficient, (c) attitude profiles, (d) space weather conditions, and (e) altitudes.

Unless otherwise stated, all the Monte Carlo simulations are run using the set-up described in [Tables 7 and 8](#). We use the same constant-step modified ‘RK-45’ integrator described in Section 8.1 for orbit propagation. Furthermore, drag coefficients are computed using machine learning models, and, similar to Section 8.1, the machine learning inputs  $T_w = 400$  K,  $T_\infty$  is from the NRLMSISE-00 density model,  $\alpha_{ads} = 1$ , and  $\alpha_{surf}$  is given by Eq. 13.

### 8.2.1. Different drag coefficient models

We consider three different drag coefficient models, the details of which are given in [Table 10](#). Cases A and B are based on sampling from a normal distribution, whereas case C is based on a first-order Gauss–Markov process ([Schutz et al., 2004](#)):

$$\kappa(t) = \exp(-\eta\Delta t)\kappa(t - \Delta t) + u_k(t) \times \sqrt{\frac{\Gamma^2}{2\eta}(1 - \exp(-2\eta\Delta t))} \quad (14a)$$

$$\eta = -\frac{\ln 0.5}{\tau} \quad (14b)$$

Table 9  
Top-Level algorithm for orbit uncertainty propagation.

---

**Algorithm:** Monte Carlo simulations

---

```

For  $j = 1 : 500$ 
   $X_0$  = initial orbital state
  For  $t = 0 : 10 \text{ seconds} : 3\text{Days}$ 
    For  $s = 1 : 6$  ( $s$  represents the six species  $H, He, N, N_2, O, O_2$ )
       $c_d(j, t, s)$  = a sampled  $C_D$  value for species  $s$  from its probability distribution function
    End loop
     $C_D(j, t)$  = total drag coefficient computed using Eq. 10 and using the sampled drag coefficients of the six species
    Using  $C_D(j, t)$ , compute the drag force. Propagate the object to the next time step
  End loop
   $X(j)$  = orbital state at the end of 3-day propagation
End loop
Compute the distribution of  $X - X_{ref}$ , where  $X_{ref}$  is some reference orbit

```

---

where  $u_k(t)$  is a random number sampled from the standard normal distribution. The factor  $(\Gamma^2/(2\eta))$ , which represents the steady-state variance of  $\kappa$ , is taken to be unity. The parameter  $\tau$  is the “half-life” and governs the rate at which the auto-correlation fades. Based on literature ([McLaughlin et al., 2012](#)), the half-life  $\tau$  is taken to be 1.8 min. At the initial epoch (i.e.,  $t = 0$ ),  $\kappa$  is simply a random number sampled from the standard normal distribution.

Case C (Gauss–Markov process based) is the most realistic in operations as drag coefficients have spatiotemporal correlations. Case A (Gaussian noise-based), on the other hand, is the least realistic because the drag coefficient distribution changes over time depending on various input factors such as atmospheric temperature, the density of atmospheric species, and others. For all the three cases A, B, and C, we assume that the attitude varies as: pitch  $= 90 \sin(100D)^\circ$  and yaw  $= 180(1 + \cos(100D))^\circ$ , where  $D$  is the time in days since the initial epoch, and the argument for sine/cosine functions are in radians. The attitude-dependent cross-section area is obtained using [LinearNDInterpolator](#).

[Fig. 8](#) shows the along-track errors at the end of three days of orbit propagation for cases A, B, and C. The reference orbit for the computation of the along-track errors is the orbit propagated with constant drag coefficient  $C$  (see [Table 10](#) for details of  $C$ ) with cross-sectional area varying according to the full-attitude profile described earlier. In [Fig. 8](#), we show both the normalized histogram and the theoretical normal probability density function (PDF) fit. Compared to case A (Gaussian noise-based), the spread in the PDF (or the standard deviation) is much larger for case C (Gauss–Markov process-based). In case A, there is little to no spatiotemporal correlation in the sampled drag coefficient values along the orbit (partial random behavior); this results in the cancellation of perturbation effects, resulting in unrealistically small orbital errors. This behavior is demonstrated in [Fig. 9](#), where we show the drag coefficients for the first five Monte Carlo samples and the reference orbit for the initial 390 s. [Table 11](#) lists the bias and the  $3\sigma$  uncertainties for the radial, along-track, and

Table 10

Description of different drag coefficient models.

Cases	$C_D$ Model (For Each Time Step Along The Orbit)
A	Total $C_D$ sampled from $\mathcal{N}(C, .01)$ , i.e., a normal distribution with constant mean $C$ and standard deviation 0.01. <i>Computation of <math>C</math></i> - Let us consider case B described below. Let an object be propagated using $\mu_{C_D}$ , which varies along the orbit. Then the constant $C$ is the average of all the drag coefficient values along that orbit.
B	For each species $i$ , 500 drag coefficient values are sampled from $\mathcal{N}(\mu_i, \sigma_i)$ . The mean $\mu_i$ and the standard deviation $\sigma_i$ come from the machine learning models for species $i$ . Total mean drag coefficient $\mu_{C_D}$ and the associated standard deviation $\sigma_{C_D}$ are computed from $\mu_i$ ( $i = 1 : 6$ ) and the 500 samples using Eq. 10.
C	Total $C_D$ sampled from $\mathcal{N}(\mu_{C_D}, \sigma_{C_D})$ Total mean drag coefficient $\mu_{C_D}$ and the associated standard deviation $\sigma_{C_D}$ are first computed in the same manner as that of case B. A parameter $\kappa$ is then sampled from a first-order Gauss–Markov process. Total drag coefficient is given as: $C_D = \mu_{C_D} + \kappa\sigma_{C_D}$

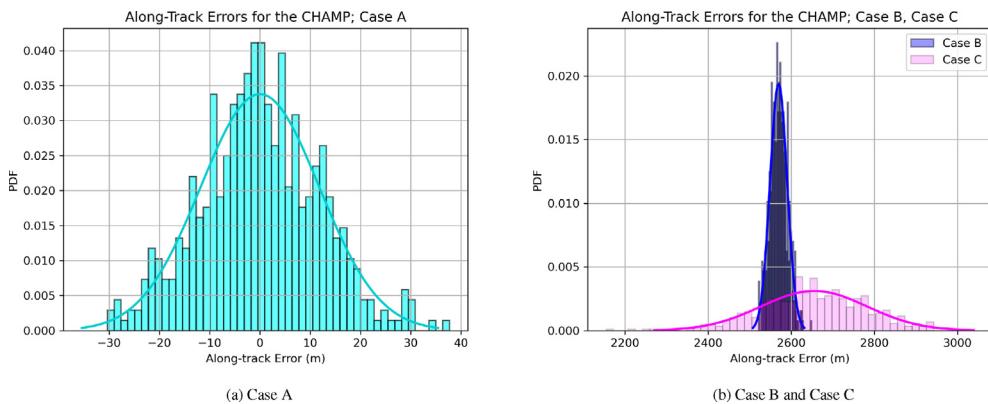


Fig. 8. Along-track errors for cases A, B, and C at the end of three days of orbit propagation. Case A: Gaussian noise-based; Case B: machine learning-based normal distribution; Case C: based on Gauss–Markov process with half-life = 1.8 min. Attitude profile: full-altitude variation. Space weather condition: geomagnetic storm, solar maximum. Study of the effect of different drag coefficient models.

cross-track errors for all three cases at the end of three days of orbit propagation. From the table, the general trend is:  $3\sigma$  for case C (Gauss–Markov process-based)  $\gg 3\sigma$  for case B (machine learning-based normal distribution)  $> 3\sigma$  for case A (Gaussian noise-based). The bias is approximately zero for case A and is of a similar order for cases B and C. Note that we do not show the plots for radial and cross-track errors because they are much smaller than the along-track errors and for brevity.

### 8.2.2. Different spatiotemporal correlations for the drag coefficient

In the first-order Gauss–Markov process-based orbit uncertainty propagation, i.e., case C (see Table 10), the half-life is taken as 1.8 min. However, in operations, the half-life can vary. We, therefore, re-run case C with a larger half-life of 18 min; we refer to this new case as ‘case D.’ Fig. 10 shows the distribution of along-track errors for cases C and D. The reference orbit we use here is the same as the one we use in Section 8.2.1. The spread (uncertainty) is much larger for case D (half-life = 18 min) because the sampled drag coefficient values are more correlated along the orbit, i.e., there is less random behavior in comparison to case C (half-life = 1.8 min). Tables 11 (for case C) and 12 (for case D) list the bias and  $3\sigma$  uncertainties for the

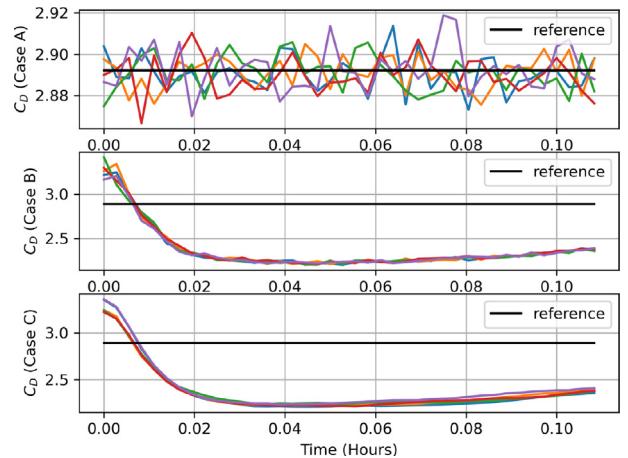


Fig. 9. Comparison of  $C_D$  samples for different drag coefficient models. Case A: Gaussian noise-based; Case B: machine learning-based normal distribution; Case C: based on Gauss–Markov process with half-life = 1.8 min. Attitude profile: full-altitude variation. Space weather condition: geomagnetic storm, solar maximum. Note that for case B and case C, the mean drag coefficient values along the orbit (not shown in the figure) would lie close to the shown Monte Carlo samples; the variability of the  $C_D$ , i.e., the difference between the  $C_D$  samples and the reference is not to be confused with the relatively small standard deviation values.

radial, along-track, and cross-track errors at the end of orbit propagation. The bias is nearly identical for cases C and D.

Table 11

Bias and  $3\sigma$  uncertainties for the radial, along-track, and cross-track errors for cases A, B, and C at the end of orbit propagation. Case A: Gaussian noise-based; Case B: machine learning-based normal distribution; Case C: based on Gauss–Markov process with half-life = 1.8 min. Attitude profile: full-attitude variation. Space weather condition: geomagnetic storm, solar maximum.

Errors	Bias (Case A   Case B   Case C)	Uncertainty ( $3\sigma$ Values) (Case A   Case B   Case C)
Radial (m)	0.0039   -4.1968   -4.5222	0.1794   0.2797   1.8238
Along-track (m)	0.0136   2569.3938   2654.6444	35.4259   61.7800   384.0162
Cross-track (m)	1.1563E-5   0.2362   0.1302	0.0034   0.0062   0.1772

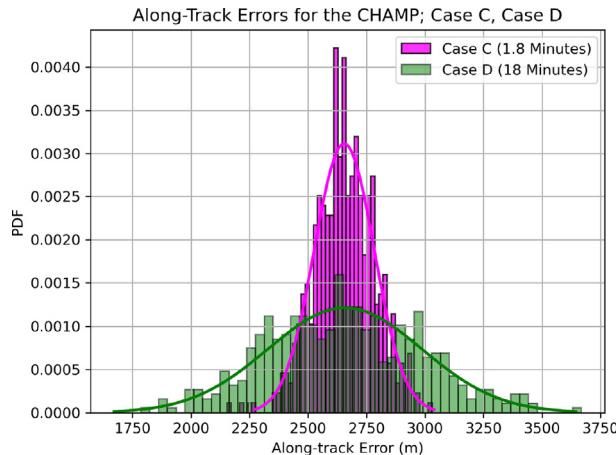


Fig. 10. Along-track errors for cases C and D at the end of three days of orbit propagation. Study of the effect of different drag coefficient spatiotemporal correlation. Case C/D: Gauss–Markov process-based. Attitude profile: full-attitude variation. Space weather condition: geomagnetic storm, solar maximum.

### 8.2.3. Different attitude profiles

For cases A, B, and C in Section 8.2.1, we assume a full-attitude variation (one can think of it as tumbling). However, space objects may have reduced-attitude profiles achieved through control mechanisms involving reaction wheels, control moment gyros, and others. We, therefore, re-run cases A, B, and C with the following attitude profile: pitch =  $\sin(100D)^\circ$ , yaw =  $5 \cos(100D)^\circ$ ; we refer to these new cases as ‘case E’, ‘case F’, and ‘case G’, respectively. Fig. 11 shows the along-track errors for cases E, F, and G at the end of orbit propagation. For Fig. 11, we generate the reference orbit using the same procedure as in Section 8.2.1, except with a reduced-attitude profile instead of the full-attitude profile. Table 13 lists the bias and  $3\sigma$  uncertainties for the radial, along-track, and cross-track errors at the end of orbit propagation for cases E, F, and

G. The  $3\sigma$  values are smaller for the reduced-attitude cases when compared to the full-attitude cases. This is because of the smaller variations in ballistic coefficient along the orbit for the reduced-attitude cases; this is demonstrated in Fig. 12, where we compare  $C_D \times Area$  for two randomly selected Monte Carlo samples – one from the full-attitude case (Case C) and one from the reduced-attitude case (case G) – for approximately six hours.

### 8.2.4. Different space weather conditions

The effect of the drag coefficient on orbital perturbations can be significantly influenced by space weather conditions (Paul et al., 2021). In the previous simulations, the propagation period roughly coincides with a geomagnetic storm during solar maximum. To study the effect of different space weather conditions, we re-run case C (Gauss–Markov process based) with an initial epoch of 00:00:00 UT, October 01, 2009, which corresponds to a quiet time during solar minimum; we refer to this new case as ‘case H.’ Fig. 13 shows the along-track errors for case H at the end of orbit propagation. The reference for Fig. 13 is an orbit generated by propagating an object with mean drag coefficients along the orbit computed from the machine learning models, full-attitude cross-sectional area variation, and quiet time space weather conditions. Table 14 lists the bias and  $3\sigma$  uncertainties for the radial, along-track, and cross-track errors at the end of propagation for case H. The uncertainty in the along-track errors for the quiet time-based case H (196.5790 m) is almost half that of the uncertainty in the along-track errors for the storm time-based case C (384.0162 m). Compared to the geomagnetic storm/solar maximum period, the atmospheric densities are much smaller during the quiet solar minimum period. There is a strong coupling between the atmospheric density and drag coefficient values, which results in the differences between cases C and H. In any mission design, it is crucial to correctly model the space weather conditions to account for the correct impact of drag coefficient uncertainties.

### 8.2.5. Different altitudes

The simulations conducted so far assume an altitude of around 400 km. To investigate the impact of satellite altitude on drag coefficient-induced orbital uncertainties, we re-run case C with an initial semi-major axis of 6628136.3000 meters (altitude  $\approx$  250 km); we refer to this new case as ‘case I.’ Since the starting position is different

Table 12

Bias and  $3\sigma$  uncertainties for the radial, along-track, and cross-track errors for case D at the end of orbit propagation. Case D: based on Gauss–Markov process with half-life = 18 min. Attitude profile: full-attitude variation. Space weather condition: geomagnetic storm, solar maximum.

Errors	Bias	Uncertainty ( $3\sigma$ Values)
Radial (m)	-4.5428	4.0219
Along-track (m)	2657.4929	986.7020
Cross-track (m)	0.1299	0.2157

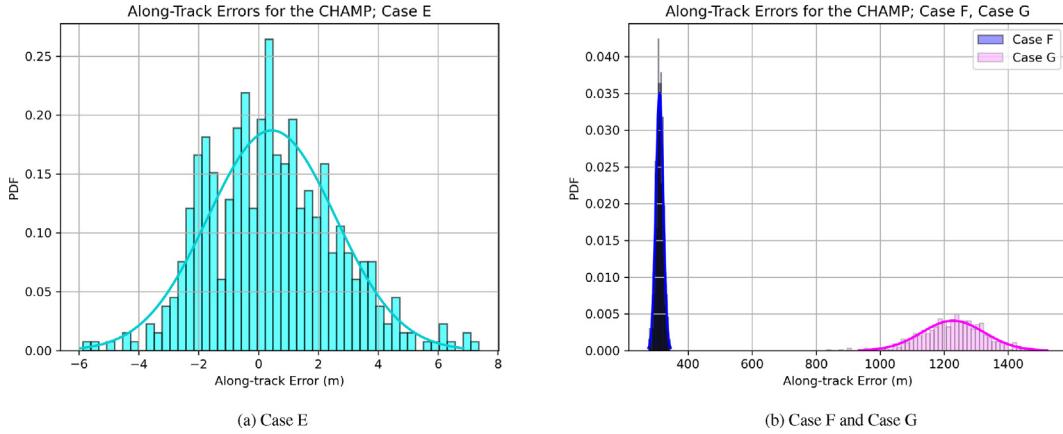


Fig. 11. Along-track errors for cases E, F, and G at the end of three days of orbit propagation. Case E: Gaussian noise-based; Case F: machine learning-based normal distribution; Case G: based on Gauss–Markov process with half-life = 1.8 min. Attitude profile: reduced-attitude variation. Space weather condition: geomagnetic storm, solar maximum. Study of the effect of different attitude profiles.

Table 13

Bias and  $3\sigma$  uncertainties for the radial, along-track, and cross-track errors for cases E, F, and G at the end of orbit propagation. Case E: Gaussian noise-based; Case F: machine learning-based normal distribution; Case G: based on Gauss–Markov process with half-life = 1.8 min. Attitude profile: reduced-attitude variation. Space weather condition: geomagnetic storm, solar maximum.

Errors	Bias (Case E   Case F   Case G)	Uncertainty ( $3\sigma$ Values) (Case E   Case F   Case G)
Radial (m)	-0.0013   -2.1898   -7.7465	0.0329   0.1642   1.5581
Along-track (m)	0.4423   310.4729   1227.5856	6.3985   34.1737   293.9133
Cross-track (m)	3.4203E-5   0.0365   0.0415	0.0007   0.0035   0.1878

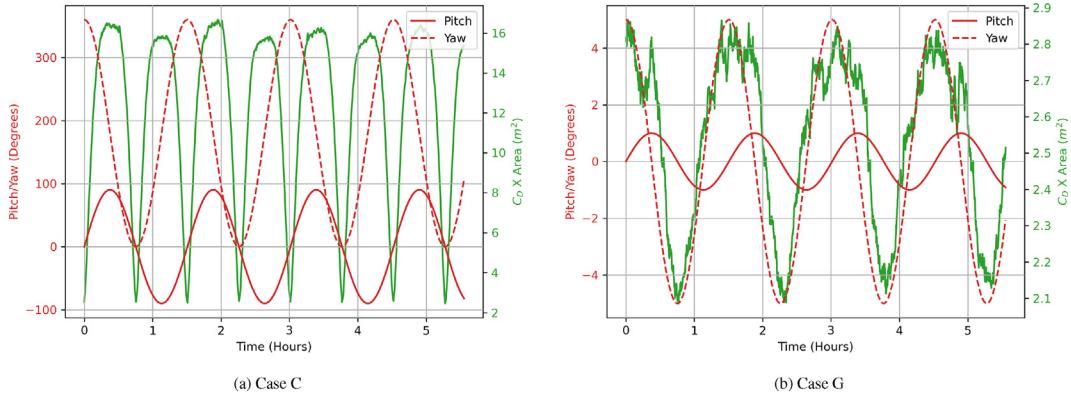


Fig. 12. Comparison of ballistic coefficient for two randomly selected Monte Carlo samples from cases C and G, respectively. Case C: based on Gauss–Markov process with half-life = 1.8 min and full-attitude variation; Case G: based on Gauss–Markov process with half-life = 1.8 min and reduced-attitude variation. Space weather condition: geomagnetic storm, solar maximum.

for case I (compared to all the previous simulation cases), a new reference orbit is defined with an initial location at an altitude of 250 km; we propagate the reference orbit with mean drag coefficients along the orbit computed from the machine learning models, full-attitude cross-sectional area variation, and storm time space weather conditions. Fig. 14 shows the along-track errors at the end of three days of orbit propagation for case I. Table 15 lists the bias and  $3\sigma$  uncertainties for the radial, along-track, and cross-track errors at the end epoch for case I. Thus, at very low altitudes, the along-track uncertainties ( $3\sigma$  values) can be of the order of 10 km, emphasizing the importance of tak-

ing drag coefficient uncertainties into account when conducting orbit propagation. Note that we did not investigate any further lower altitudes; at altitudes below 200 km, the free molecular flow assumption starts to break down, and our machine learning models are not valid in the continuum flow regime.

## 9. Application and limitations

We intend to add the proposed drag coefficient modeling approach to the WVU RSM toolkit to help satellite owners/operators with their spacecraft operations and life-

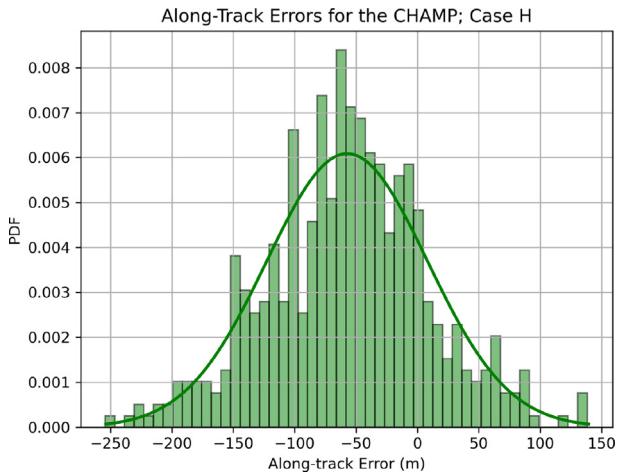


Fig. 13. Along-track errors for case H at the end of three days of orbit propagation. Case H: based on Gauss–Markov process with half-life = 1.8 min. Attitude profile: full-attitude variation. Space weather condition: quiet, solar minimum.

Table 14

Bias and  $3\sigma$  uncertainties for the radial, along-track, and cross-track errors for case H at the end of orbit propagation. Case H: based on Gauss–Markov process with half-life = 1.8 min. Attitude profile: full-attitude variation. Space weather condition: quiet, solar minimum.

Errors	Bias	Uncertainty ( $3\sigma$ Values))
Radial (m)	0.3740	1.1161
Along-track (m)	-57.2505	196.5790
Cross-track (m)	0.0501	0.1566

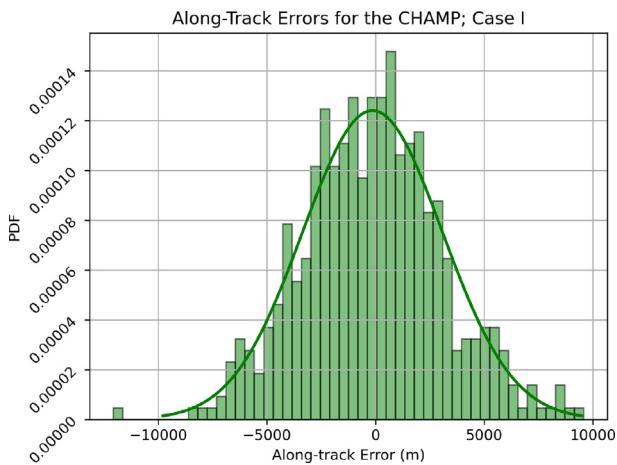


Fig. 14. Along-track errors for case I at the end of three days of orbit propagation. Case I: based on Gauss–Markov process with half-life = 1.8 min. Attitude profile: full-attitude variation. Altitude: approximately 250 km. Space weather condition: geomagnetic storm, solar maximum.

time prediction. Particularly, for operations involving large attitude changes, such as the application of differential drag, or specific science mission maneuvers (e.g., the GRACE satellite(s) did 180° yaw turns), our drag coefficient modeling approach can be applied. At the end of their

Table 15

Bias and  $3\sigma$  uncertainties for the radial, along-track, and cross-track errors for case I at the end of orbit propagation. Case I: based on Gauss–Markov process with half-life = 1.8 min. Attitude profile: full-attitude variation. Altitude: approximately 250 km. Space weather condition: geomagnetic storm, solar maximum.

Errors	Bias	Uncertainty ( $3\sigma$ Values))
Radial (m)	1.3790	65.5767
Along-track (m)	-140.2569	9651.4373
Cross-track (m)	0.1898	1.1521

operational life, many of the current and upcoming satellites will become defunct, possibly with a random tumbling profile. With a broader push in the space community now for owner-operators to release information on defunct satellites' geometry and surface properties, our methodology (which requires sufficient fidelity geometry and surface property know-how) can be used to predict the trajectories of these objects. There is also a science application for our approach - estimates of atmospheric density can be obtained from good estimates of drag coefficient and satellite acceleration measurements. A point to note is that the developed approach can only be applied to space debris or classified objects if the geometry and surface properties can be estimated. For space debris, we may not readily know the object characteristics, so we will have to use photometric/astrometric data to determine the object characteristics before carrying out the numerical simulations to generate the training data for the machine learning algorithms.

## 10. Conclusions and future possibilities

In this paper, we use stochastic machine learning techniques to design surrogate models for predicting the physical drag coefficient for the complex CHAMP satellite. In our study, we demonstrate that the machine learning models, which are computationally much faster than the numerical methods, are able to emulate drag coefficient results with sufficient accuracy.

Unlike other surrogate models, such as the Gaussian Process regression, we particularly stress the computational efficiency, model accuracy, reliability of the predicted uncertainties, reusability of models, and applicability of the models to all attitudes in this study. We develop a methodology to determine the appropriate data size for training the feed-forward deep neural network models. With approximately 50,000 data points for training and another 50,000 data points for testing purposes, our models are able to predict drag coefficients for  $H, He, N, N_2, O, O_2$  with root mean squared errors of 0.0295, 0.0188, 0.0250, 0.0159, 0.0142, 0.0209 and mean absolute calibration errors of 0.8845%, 2.2404%, 1.0279%, 1.3170%, 0.5117%, and 0.8384%, respectively.

To establish the importance of drag coefficient uncertainties for space operations and space situational awareness purposes, we carry out orbit uncertainty propagation via Monte Carlo simulations for a variety of

(a) drag coefficient models, (b) spatiotemporal correlations for the drag coefficient, (c) attitude profiles, (d) space weather conditions, and (e) altitudes. In orbit propagation, uncertainty in the drag coefficient mainly manifests itself in the form of uncertainty in along-track errors. We demonstrate a number of important observations - (1) uncertainty in along-track errors can be an order of magnitude higher for realistic sampling methods such as the first-order Gauss Markov process when compared to the naive sampling from a normal distribution, (2) the selection of correct values of parameters modeling the drag coefficient spatiotemporal correlation can significantly change the distribution of the orbital uncertainties, (3) orbital uncertainties are strongly affected by the attitude profile of the space object, (d) the effect of the drag coefficient uncertainties on orbital state uncertainties is much more significant under a geomagnetic storm during solar maximum when compared to a quiet period during solar minimum, and (e) the uncertainty in the along-track errors because of drag coefficient uncertainties can be of the order of 10 km (or more) for a low altitude of 250 km.

The current study only considers one gas-surface interaction model. In the future, we intend to explore different kinds of gas-surface interaction, which is the largest source of uncertainty in drag coefficient modeling. Additionally, note that the investigations conducted in this paper only concern themselves with the uncertainties emanating from fitting the stochastic models to the training data. In the future, we will also include the effects of the uncertainties in inputs, such as the atmospheric composition, temperature, and others resulting from the space weather uncertainties.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors would like to acknowledge the use of the High Performance Computing Resources Thorny Flat at West Virginia University housed at the Pittsburgh Supercomputing Center, funded by the National Science Foundation (NSF) Major Research Instrumentation Program (MRI) Award #1726534.

## References

- Anderson, G.J., Gaffney, J.A., Spears, B.K., Bremer, P.-T., Anirudh, R., Thiagarajan, J.J., 2020. Meaningful uncertainties from deep neural network surrogates of large-scale numerical simulations. <https://doi.org/10.48550/ARXIV.2010.13749>.
- Camporeale, E., Carè, A., 2020. Estimation of accurate and calibrated uncertainties in deterministic models. <https://doi.org/10.48550/ARXIV.2003.05103>.

- Chambre, P.A., Schaaf, S.A., 1961. Flow of Rarefied Gases. Princeton University Press, Princeton, NJ.
- Chollet, F. et al., 2015. Keras. <https://keras.io>.
- Davis, D.H., 1960. Monte Carlo calculation of molecular flow rates through a cylindrical elbow and pipes of other shapes. *J. Appl. Phys.* 31 (7), 1169–1176. <https://doi.org/10.1063/1.1735797>.
- Dormand, J., Prince, P., 1980. A family of embedded Runge-Kutta formulae. *J. Comput. Appl. Math.* 6 (1), 19–26. [https://doi.org/10.1016/0771-050X\(80\)90013-3](https://doi.org/10.1016/0771-050X(80)90013-3).
- Drob, D.P., Emmert, J.T., Crowley, G., Picone, J.M., Shepherd, G.G., Skinner, W., Hays, P., Niciejewski, R.J., Larsen, M., She, C.Y., Meriwether, J.W., Hernandez, G., Jarvis, M.J., Sipler, D.P., Tepley, C. A., O'Brien, M.S., Bowman, J.R., Wu, Q., Murayama, Y., Kawamura, S., Reid, I.M., Vincent, R.A., 2008. An empirical model of the Earth's horizontal wind fields: HWM07. *J. Geophys. Res.: Space Phys.* 113 (A12). <https://doi.org/10.1029/2008JA013668>.
- Gal, Y., Ghahramani, Z., 2016. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In: Balcan, M.F., Weinberger, K.Q. (Eds.), Proceedings of The 33rd International Conference on Machine Learning, New York, New York, USA: PMLR volume 48 of Proceedings of Machine Learning Research, pp. 1050–1059.
- Knudsen, M., 1916. Das Cosinusgesetz in der kinetischen Gastheorie. *Ann. Phys.* 353 (24), 1113–1121. <https://doi.org/10.1002/andp.19163532409>.
- Laves, M.-H., Ihler, S., Fast, J.F., Kahrs, L.A., Ortmaier, T., 2021. Recalibration of aleatoric and epistemic regression uncertainty in medical imaging. *Machine Learn. Biomed. Imag.* 1, URL: <https://melba-journal.org/papers/2021:008.html>.
- Lemmens, S., Letizia, F., 2020. Space traffic management through environment capacity. *Handbook of Space Security: Policies, Applications and Programs*, pp. 845–864.
- Licata, R.J., Mehta, P.M., 2022. Uncertainty quantification techniques for data-driven space weather modeling: thermospheric density application. *Sci. Rep.* 12 (1). <https://doi.org/10.1038/s41598-022-11049-3>.
- March, G., IJssel, J.V.d., Siemes, C., Visser, P.N.A.M., Doornbos, E.N., Pilinski, M., 2021. Gas-surface interactions modelling influence on satellite aerodynamics and thermosphere mass density. *J. Space Weather Space Climate* 11 (54). <https://doi.org/10.1051/swsc/2021035>.
- March, G., Visser, T., Visser, P., Doornbos, E., 2019. CHAMP and GOCE thermospheric wind characterization with improved gas-surface interactions modelling. *Adv. Space Res.* 64 (6), 1225–1242. <https://doi.org/10.1016/j.asr.2019.06.023>.
- Matzka, J., Bronkalla, O., Tornow, K., Elger, K., Stolle, C., 2021. Geomagnetic Kp index. V. 1.0. GFZ Data Services, <https://doi.org/10.5880/Kp.0001>.
- McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21 (2), 239–245. <https://doi.org/10.1080/00401706.1979.10489755>.
- McLaughlin, C.A., Lechtenberg, T., Fattig, E., Krishna, D.M., 2012. Estimating density using precision satellite orbits from multiple satellites. *J. Astronaut. Sci.* 59, 84–100. <https://doi.org/10.1007/s40295-013-0007-4>.
- Mehta, P.M., Paul, S.N., Crisp, N.H., Sheridan, P.L., Siemes, C., March, G., Bruinsma, S., 2022. Satellite drag coefficient modeling for thermosphere science and mission operations. *Adv. Space Res.* <https://doi.org/10.1016/j.asr.2022.05.064>.
- Mehta, P.M., Walker, A., Lawrence, E., Linares, R., Higdon, D., Koller, J., 2014a. Modeling satellite drag coefficients with response surfaces. *Adv. Space Res.* 54 (8), 1590–1607. <https://doi.org/10.1016/j.asr.2014.06.033>, URL: <https://www.sciencedirect.com/science/article/pii/S027311771400413X>.
- Mehta, P.M., Walker, A., McLaughlin, C.A., Koller, J., 2014b. Comparing physical drag coefficients computed using different gas-surface interaction models. *J. Spacecraft Rock.* 51 (3), 873–883. <https://doi.org/10.2514/1.A32566>.

- Moe, K., Moe, M.M., Rice, C.J., 2004. Simultaneous analysis of multi-instrument satellite measurements of atmospheric density. *J. Spacecraft Rock.* 41 (5), 849–853. <https://doi.org/10.2514/6.2003-570>.
- Mostaza Prieto, D., Graziano, B.P., Roberts, P.C., 2014. Spacecraft drag modelling. *Prog. Aerosp. Sci.* 64, 56–65. <https://doi.org/10.1016/j.paerosci.2013.09.001>.
- O'Malley, T., Bursztein, E., Long, J., Chollet, F., Jin, H., Invernizzi, L., et al., 2019. Keras tuner. <https://github.com/keras-team/keras-tuner>.
- Paul, S., Sheridan, P., Mehta, P., Huzurbazar, S., 2021. Satellite drag coefficient modeling and orbit uncertainty quantification using stochastic machine learning techniques. In: in Proc. 2021 AAS/AIAA Astrodynamics Specialist Conf.
- Picone, J.M., Hedin, A.E., Drob, D.P., Aikin, A.C., 2002. NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *J. Geophys. Res.: Space Phys.* 107 (A12), SIA 15-1–SIA 15-16. <https://doi.org/10.1029/2002JA009430>.
- Pilinski, M.D., Argrow, B.M., Palo, S.E., 2010. Semi-empirical model for satellite energy-accommodation coefficients. *J. Spacecraft Rock.* 47 (6), 951–956. <https://doi.org/10.2514/1.49330>.
- Plimpton, S., Moore, S., Borner, A., Stagg, A., Koehler, T., Torczynski, J., Gallis, M., 2019. Direct simulation Monte Carlo on petaflop supercomputers and beyond. *Phys. Fluids* 31 (8), 86–101.
- Raissi, M., Babaei, H., Karniadakis, G.E., 2019. Parametric Gaussian Process regression for big data. *Comput. Mech.* 64, 409–416. <https://doi.org/10.1007/s00466-019-01711-5>.
- Rasmussen, C.E., 2004. Gaussian processes in machine learning. In: Bousquet, O., von Luxburg, U., Rätsch, G. (Eds.), Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2–14, 2003, Tübingen, Germany, August 4–16, 2003, Revised Lectures. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 63–71. [https://doi.org/10.1007/978-3-540-28650-9\\_4](https://doi.org/10.1007/978-3-540-28650-9_4).
- Schutz, B., Tapley, B., Born, G., 2004. Statistical Orbit Determination. chapter 6, 1st ed. Elsevier Academic Press, Burlington, MA.
- Sheridan, P.L., Paul, S.N., Avendaño-Franco, G., Mehta, P.M., 2022. Updates and improvements to the satellite drag coefficient response surface modeling toolkit. *Adv. Space Res.* 69 (10), 3828–3846. <https://doi.org/10.1016/j.asr.2022.02.044>.
- Sun, F., Gramacy, R.B., Haaland, B., Lawrence, E., Walker, A., 2019a. Emulating satellite drag from large simulation experiments. SIAM/ASA J. Uncert. Quantif. 7 (2), 720–759. <https://doi.org/10.1137/18M1170157>.
- Sun, F., Gramacy, R.B., Haaland, B., Lawrence, E., Walker, A., 2019b. Emulating satellite drag from large simulation experiments. SIAM/ASA J. Uncert. Quantif. 7 (2), 720–759. <https://doi.org/10.1137/18M1170157>.
- Sutton, E.K., 2009. Normalized force coefficients for satellites with elongated shapes. *J. Spacecraft Rock.* 46 (1), 112–116. <https://doi.org/10.2514/1.40940>.
- Svozil, D., Kvasnicka, V., Pospichal, J., 1997. Introduction to multi-layer feed-forward neural networks. *Chemomet. Intell. Lab. Syst.* 39 (1), 43–62. [https://doi.org/10.1016/S0169-7439\(97\)00061-0](https://doi.org/10.1016/S0169-7439(97)00061-0).
- Tapping, K.F., 2013. The 10.7 cm solar radio flux (f10.7). *Space Weather* 11 (7), 394–406. <https://doi.org/10.1002/swe.20064>.
- van der Wilk, M., Dutordoir, V., John, S., Artemev, A., Adam, V., Hensman, J., 2020. A framework for interdomain and multioutput Gaussian Processes. arXiv:2003.01115, URL: <https://arxiv.org/abs/2003.01115>.
- Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, I., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P. SciPy 1.0 Contributors, 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* 17, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>.
- Walker, A., Mehta, P., Koller, J., 2014a. Different implementations of diffuse reflection with incomplete accommodation for drag coefficient modeling. *J. Spacecraft Rock.* 51 (5), 1522–1532. <https://doi.org/10.2514/1.A32668>.
- Walker, A., Mehta, P., Koller, J., 2014b. Drag coefficient model using the Cercignani–Lampis–Lord gas–surface interaction model. *J. Spacecraft Rock.* 51 (5), 1544–1563. <https://doi.org/10.2514/1.A32677>.
- Walker, A., Mehta, P., Koller, J., 2014c. Drag coefficient model using the Cercignani–Lampis–Lord gas–surface interaction model. *J. Spacecraft Rock.* 51 (5), 1544–1563.