Multi-Model Monte Carlo Estimators for Trajectory Simulation

James E. Warner*, Geoffrey F. Bomarito**, Luke Morrill†, Patrick E. Leser**, William P. Leser**, Robert A. Williams‡, and Soumyo Dutta§
NASA Langley Research Center, Hampton, VA 23681, USA

Samantha Nieomoeller¶
University of California - Los Angeles, Los Angeles, CA 90095, USA

Predicting landing radius and other quantities of interest (QoI) for entry, descent, and landing (EDL) applications requires a viable uncertainty propagation method for quantifying the impact of uncertainties in aerodynamics, atmosphere, mass properties, etc. While standard Monte Carlo (MC) simulation is the de facto standard for producing robust and unbiased statistical estimators, it is often infeasible for expensive, high-fidelity models. Low-fidelity models are commonly constructed to replace the high-fidelity model in MC simulation for computational speedup, but at the expense of accuracy and unbiasedness. Emerging multi-model MC methods are bridging this gap by combining predictions from two or more models of varying fidelity and computational cost for efficient and unbiased uncertainty propagation. This works establishes a proof of concept for using multi-model MC to increase the speed and precision of trajectory simulation for EDL. It is shown that combining a high-fidelity EDL model with low-fidelity models (e.g., data-driven, reduced physics) in this manner has the potential to yield significant efficiency and accuracy gains for certain EDL QoIs versus a standard MC approach. Moreover, the unbiasedness of multi-model MC predictions is highlighted by showing increased accuracy versus an approach that leverages a low-fidelity model alone.

I. Introduction

NASA's aspirations for human class Mars missions are contingent upon significant breakthroughs in entry, descent, and landing (EDL) systems. To support a crew of four on the Martian surface, NASA's Evolvable Mars Campaign (EMC) has considered landing four separate 20 metric ton payloads that must be delivered within a 50 m radius of the desired target [1]. In order to accomplish this feat, NASA needs an order-of-magnitude increase in payload capability and nearly three orders-of-magnitude reduction in landing radius over the current state-of-the-art [2]. This work begins to address the latter challenge by exploring new methods for precise trajectory simulation.

Trajectory simulation is complicated by uncertainties in a range of factors that impact a quantity of interest (QoI), including atmospheric variations, aerodynamic uncertainties, mass properties, and initial condition variations. Predicting a vehicle's landing radius or other QoIs requires simulating thousands or even millions of potential trajectories by varying these uncertainties using a Monte Carlo (MC) dispersion analysis [2, 3]. Since time-consuming, high-fidelity EDL models are needed for accurate trajectory simulation, it can be challenging to deliver precise predictions for missions with strict time constraints.

Broadly speaking, the problem of estimating the statistics of model QoIs given uncertainties in input parameters is that of uncertainty propagation. While MC simulation can often be computationally prohibitive for expensive models, it is still the most well-known and general-purpose approach for uncertainty propagation. MC estimators are known to be robust, unbiased, and have a convergence rate that is independent of dimensionality (i.e., the number of random model inputs). A common approach to reduce the computational expense is to construct low-fidelity surrogate models to replace the high-fidelity model for MC simulation. While substantial speedup can be obtained, relying solely on the low-fidelity model for prediction will generally lead to biased and potentially inaccurate estimators [4].

*Research Computer Scientist, Durability, Damage Tolerance, and Reliability Branch, M/S 188E.
**Materials Engineer, Durability, Damage Tolerance, and Reliability Branch, M/S 188E.
†Data Systems Engineer, Climate Science Branch, M/S 475
‡Research Computer Scientist, Atmospheric Flight and Entry Systems Branch, M/S 489.
§Aerospace Engineer, Atmospheric Flight and Entry Systems Branch, M/S 489, AIAA Senior Member.
¶Graduate Student, Computer Science Department

Multi-model MC methods have recently emerged to bridge this gap, combining information from two or more models of varying fidelity and computational cost to produce efficient, accurate predictions. By retaining the high-fidelity model, unbiased estimators of a QoI can be constructed, while introducing an ensemble of low-fidelity models provides computational speedup. Existing multi-model methods are distinguished by the types of models used and the approach to allocate computational resources across the available models. Multilevel MC (MLMC) [5] restricts low-fidelity models to those arising from coarsened discretizations of the same governing equations in space/time, while multifidelity MC (MFMC) [6] generalizes the types of models permitted to have different forms (e.g., data-driven, reduced-order, analytical, etc.). Both MLMC and MFMC introduce simplifying assumptions that yield closed form expressions for resource allocation across models. A generalized approximate control variates (ACVs) [7] approach was later proposed and then further developed [8] that unified and improved upon MLMC-based and MFMC methods. It was shown that significant performance improvements could be obtained relative to MLMC and MFMC by optimizing for resource allocation with a more complex numerical approach.

This works establishes a proof of concept for using multi-model MC to increase the speed and precision of trajectory simulation for EDL. By leveraging recently-released NASA software [9] that implements MLMC, MFMC, ACV, and other multi-model methods, estimators for relevant EDL QoIs were constructed and assessed with respect to standard MC simulation. Program to Optimize Simulated Trajectories II (POST2) [10], a legacy code developed at NASA Langley Research Center, was used for high-fidelity trajectory simulation. Several low-fidelity models were constructed for use within the multi-model framework, including a data-driven machine learning model trained from POST2 input/output data, as well as lower-fidelity POST2 simulations that used significantly larger time steps and simplified physics modules. As an initial demonstration of the approach, the focus here is limited to estimating the expected value of QoIs, providing a foundation for extensions of multi-model MC that recover higher order statistics [15, 16] and probability density functions [17] in future studies.

An analysis of the Adaptable Deployable Entry and Placement Technology (ADEPT) Sounding Rocket 1 (SR-1) test flight [3] was revisited to quantify accuracy and computational cost for predicting several relevant EDL QoIs. It is demonstrated that a multi-model MC approach can substantially outperform standard MC simulation that relies on either POST2 or a low-fidelity surrogate model alone. In particular, multi-model MC achieves 1-2 orders of magnitude reduction in mean squared error (MSE) for QoIs such as landing location, altitudes, and roll rates, while in the worst case, the performance is comparable to standard MC for more dynamic QoIs, such as the angle of attack along the trajectory. This variation in performance is shown to result from an insufficient degree of correlation between the particular low-fidelity models considered relative to the high-fidelity model for the dynamic QoIs, motivating future research on model development.

The paper is organized as follows. Section II provides the mathematical background for the work, introducing trajectory simulation as an uncertainty propagation problem before describing both the standard MC and multi-model MC solutions. Section III gives the details of the EDL application that is the focus of this work, including an overview of the ADEPT test flight as well as a description of the various trajectory simulation models considered within the multi-model MC approach. Results comparing the performance of standard and multi-model MC for predicting ADEPT QoIs are then illustrated in Section IV. Finally, Section V provides conclusions of this study and discusses potential future study.

II. Background

A. Uncertainty Propagation & Monte Carlo Estimators

Predicting landing location and other QoIs for EDL under uncertainty in parameters governing atmospheric variations, initial vehicle conditions, etc. can be described mathematically as an uncertainty propagation problem. Let $Y_0 \in \mathbb{R}^1$ represent a scalar QoI and $\mathcal{M}_0$ define a computational model that maps some input parameters, $\mathbf{X} \in \mathbb{R}^d$, to the QoI:

$$\mathcal{M}_0 : \mathbf{X} \rightarrow Y_0. \tag{1}$$

Then, given a description of uncertainty in $\mathbf{X}$ (i.e., its probability density function, $p(\mathbf{x})$), uncertainty propagation methods seek to estimate the statistics of Y_0 using $\mathcal{M}_0$.

In particular, this work focuses primarily on the problem of constructing accurate estimators for the expected value of a QoI,

$$\mathbb{E}[Y_0] = \int Y_0(\mathbf{x}) p(\mathbf{x}) d\mathbf{x}, \tag{2}$$

when $\mathcal{M}_0$ is an expensive, high-fidelity model. Using standard MC simulation, this integral can be approximated by an estimator for $\mathbb{E}[Y_0]$ of the following form:

$$\hat{Y}_0 = \frac{1}{N_0} \sum_{i=1}^{N_0} Y_0(\mathbf{x}^{(i)}), \tag{3}$$

where $\{\mathbf{x}^{(i)}\}_{i=1}^{N_0}$ is a collection of independent and identically distributed (i.i.d.) samples from $p(\mathbf{x})$. Note that $\hat{Y}_0 \equiv \hat{Y}_0(\{\mathbf{x}^{(i)}\}_{i=1}^{N_0})$ is a random variable itself since it is dependent on a randomly drawn set of input samples.

The mean squared error (MSE) associated with a random estimator $\hat{Y}_0$ for $\mathbb{E}[Y_0]$ can be decomposed into bias and variance terms as follows

$$e(\hat{Y}_0) \equiv \mathbb{E}\left[(\hat{Y}_0 - \mathbb{E}[Y_0])^2\right] = \left(\text{Bias}[\hat{Y}_0]\right)^2 + \text{Var}[\hat{Y}_0], \tag{4}$$

where

$$\text{Bias}[\hat{Y}_0] = \mathbb{E}[\hat{Y}_0] - \mathbb{E}[Y_0] \tag{5}$$

and

$$\text{Var}[\hat{Y}_0] = \mathbb{E}\left[(\mathbb{E}[\hat{Y}_0] - \mathbb{E}[Y_0])^2\right]. \tag{6}$$

Since the estimator in Equation (3) for standard MC is unbiased, $\text{Bias}[\hat{Y}_0] = 0$ in the equations above and the MSE is equivalent to the MC estimator variance [11], given by

$$\text{Var}[\hat{Y}_0] = \frac{\text{Var}[Y_0]}{N_0}. \tag{7}$$

Equation (7) shows that the convergence rate of standard MC estimators is relatively slow ($O(N_0^{-1})$) and so a huge number of model evaluations can be required to obtain accurate estimators, especially when $\text{Var}[Y_0]$ is large. Thus, using MC simulation when it is time consuming to evaluate $\mathcal{M}_0$ can be impractical. Formally, if the cost (run time) to evaluate $\mathcal{M}_0$ is C_0, then the total cost to generate the MC estimator in Equation (3) is

$$\hat{C} = N_0 C_0. \tag{8}$$

Although high performance computing can be utilized to perform the N_0 model evaluations in parallel, this total cost can still be prohibitive for EDL applications with strict time constraints. Note that Equation (8) is an idealized total run time that assumes additional costs from operating system overheads, evaluating Equation (3), etc. are negligible relative to the cost of one high fidelity simulation.

When the computational cost of standard MC is intractable, a common strategy is to introduce a low-fidelity model *, $\mathcal{M}_1 : \mathbf{X} \to Y_1$, that approximates the high fidelity model in Equation (1) but is less expensive to evaluate ($C_1 \ll C_0$)[12]. Then, $\mathcal{M}_1$ can be used *instead of* $\mathcal{M}_0$ to form an estimator, $\hat{Y}_1$, with Equation (3). The low-fidelity estimator can result in significantly lower estimator variance (Equation (7)) since a larger number of samples, $N_1 \gg N_0$, can be used for the same total cost. However, the downside of this approach is that $\hat{Y}_1$ is not an unbiased estimator of $\mathbb{E}[Y_0]$ since in general $\mathcal{M}_1(\mathbf{x}^*) \neq \mathcal{M}_0(\mathbf{x}^*)$ for an arbitrary input sample, $\mathbf{x}^*$. Note that the estimator bias, $\text{Bias}[\hat{Y}_1]$, is an irreducible error that cannot be eliminated by increasing the number of samples used and so $e(\hat{Y}_1) \not\to 0$ as $N_1 \to \infty$.

B. The Multi-Model Monte Carlo Approach

This work adopts the multi-model MC approach to construct estimators for $\mathbb{E}[Y_0]$ that are both unbiased like high-fidelity MC *and* efficient like low-fidelity MC. Here, it is assumed that there exists a collection of $M \geq 1$ low-fidelity models

$$\mathcal{M}_i : \mathbf{X} \to Y_i, \quad i = 1, ..., M, \tag{9}$$

that provide faster, but potentially less accurate, approximations of Y_0 predicted by the high-fidelity model, $\mathcal{M}_0$. A variety of forms of $\mathcal{M}_i$ are possible, including simple analytical models, data-driven machine learning models, or coarser discretizations of the same physics if, for example, $\mathcal{M}_0$ represents a model derived from discretized differential equations. As long as a low-fidelity model exhibits faster run time (i.e., lower cost) than the high-fidelity model and displays some degree of correlation with respect to Y_0, it can potentially increase the performance of a multi-model MC approach.

*Depending on how it is constructed and the context in which it is applied, a low-fidelity model may be referred to as a *surrogate model*, *metamodel*, *response surface*, or *reduced-order model* [12].

A multi-model MC estimator for $\mathbb{E}[Y_0]$ can be defined as follows

$$\tilde{Y}_0(\boldsymbol{\alpha}, \{\mathbf{x}\}) = \hat{Y}_0(\mathbf{x}_0) + \sum_{i=1}^{M} \alpha_i \left(\hat{Y}_i(\mathbf{x}_{i+}) - \hat{Y}_i(\mathbf{x}_{i-}) \right), \tag{10}$$

where $\boldsymbol{\alpha}$ is a set of weighting coefficients and $\{\mathbf{x}\} \equiv \{\mathbf{x}_0, \mathbf{x}_{1+}, \mathbf{x}_{1-}, ..., \mathbf{x}_{M+}, \mathbf{x}_{M-}\}$ is a collection of potentially overlapping sets of random input samples, hereinafter referred to as the *sample allocation* [7]. Here, the inputs for each low-fidelity model are partitioned into two subsets, $\mathbf{x}_{i+} \in \mathbf{x}_i$ and $\mathbf{x}_{i-} \in \mathbf{x}_i$, such that $\mathbf{x}_{i+} \cup \mathbf{x}_{i-} = \mathbf{x}_i$ and $\mathbf{x}_{i+} \cap \mathbf{x}_{i-} \neq \emptyset$ (i.e., the subsets generally overlap). Equation (10) is a generalization of the control variates method [13, 14], a classical variance reduction technique, for cases where Y_i are random variables with unknown expected values and must be estimated from the samples $\mathbf{x}_{i-}$. Since $\hat{Y}_0$ is an unbiased MC estimator and the bias of the $\hat{Y}_i$ terms cancel out, it follows that $\tilde{Y}_0$ is also unbiased.

Let N_i denote the total number of evaluations of model $\mathcal{M}_i$ required to construct the estimator (10),

$$N_i = |\mathbf{x}_{i+} \cup \mathbf{x}_{i-}|, \tag{11}$$

where $|\cdot|$ is the cardinality of the set. Furthermore, let C_i be the cost of $\mathcal{M}_i$. Then the total cost associated with the multi-model MC estimator is

$$\tilde{C} = N_0 C_0 + \sum_{i=1}^{M} N_i C_i. \tag{12}$$

The key to multi-model MC approaches is the optimal allocation of computational resources across the available models, where *optimal* is defined as yielding an estimator in Equation (10) with the lowest MSE for a given computational budget, C^*. Knowing that $\tilde{Y}_0$ is unbiased and $e(\tilde{Y}_0) = \mathrm{Var}[\tilde{Y}_0]$, the coefficients and sample allocation for the optimal multi-model estimator can be obtained through the optimization problem

$$\begin{aligned} \boldsymbol{\alpha}^{opt}, \{\mathbf{x}\}^{opt} &= \min_{\boldsymbol{\alpha}, \{\mathbf{x}\}} \mathrm{Var}[\tilde{Y}_0(\boldsymbol{\alpha}, \{\mathbf{x}\})] \\ &\text{s.t. } \tilde{C} \leq C^*, \end{aligned} \tag{13}$$

which yields the following optimal estimator via Equation (10):

$$\tilde{Y}_0^{opt} = \tilde{Y}_0(\boldsymbol{\alpha}^{opt}, \{\mathbf{x}\}^{opt}), \tag{14}$$

with corresponding variance

$$\tilde{V}^{opt} = \mathrm{Var}[\tilde{Y}_0^{opt}]. \tag{15}$$

It is important to note that the optimization in (13) involves determining the best way to allocate a set of samples to the available models. The actual values of the samples are not optimized and are instead drawn from the input parameter distribution, $p(\mathbf{x})$. This detail was omitted for notational convenience.

1. Practical Considerations

While the details of the solution of Equation (13) are omitted here for brevity, a few comments are provided related to its implementation in practice. First, in order to evaluate $\mathrm{Var}[\tilde{Y}_0]$ during the course of the optimization solution, the covariance matrix between the model QoIs

$$\mathrm{Cov}[Y_i, Y_j] = \mathbb{E}[(Y_i - \mathbb{E}[Y_i])(Y_j - \mathbb{E}[Y_j])] \tag{16}$$

must be prescribed or estimated beforehand by executing the models for a number of *pilot samples*. The costs associated with each model, C_i, must also be estimated to enforce the computational budget constraint. Practically speaking, the solution of Equation (13) typically leads to relatively few high fidelity model evaluations and many evaluations of the low-fidelity models, especially the models that are highly correlated and computationally efficient with respect to the high-fidelity model. Second, since specifying the sample allocation, $\{\mathbf{x}\}$, includes not only the total number of samples for each model, N_i, but also their relative division into $2M + 1$ subsets, the search space for the optimization problem is intractably large. Therefore, a range of multi-model MC methods have emerged that vary on the assumptions made and strategy employed to solve Equation (13) in an efficient manner [5–8]. Since this work focuses on the practical aspects of applying these methods, the interested reader is referred to the references for more details about multi-model MC theory.

2. Notable Multi-model Monte Carlo Extensions

While the preceding formulation focused on constructing estimators for the expected value of a scalar QoI, there are several practical generalizations of multi-model MC theory worth mentioning. First, vector-valued QoIs, $\mathbf{Y}_0 = [Y_0^{(1)}, \ldots, Y_0^{(d')}] \in \mathbb{R}^{d'}$, can be easily accommodated through the modification of Equation (13) and will be demonstrated in the EDL application to follow. Here, the optimal estimator coefficients and sample allocation can be found by minimizing a weighted sum of estimator variances of each individual QoI;

$$\{\boldsymbol{\alpha}\}^{opt}, \{\mathbf{x}\}^{opt} = \min_{\{\boldsymbol{\alpha}\},\{\mathbf{x}\}} \sum_{i=1}^{d'} \gamma_i \mathrm{Var}[\tilde{Y}_0^{(i)}(\boldsymbol{\alpha}^{(i)}, \{\mathbf{x}\})] \\ \text{s.t. } \tilde{C} \leq C^*, \tag{17}$$

where γ_i are weighting coefficients adjusted to reflect relative importance and each QoI has a corresponding $\boldsymbol{\alpha}^{(i)}$ such that $\{\boldsymbol{\alpha}\} = \{\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(d')}\}$. While beyond the scope of the current study, it is also worth noting that the estimation of higher order statistics and probability density functions using multi-model MC approaches have been previously demonstrated as well [15–17].

Finally, in practice, a number of models of varying type and fidelity might be available for constructing the multi-model MC estimator. But, in general, the multi-model MC formulation assumes that all models must have samples allocated to them, even if some models may be detrimental to the overall performance of the estimator. It is difficult to identify *a priori* what subset of models will produce an estimator with minimal variance based solely on the estimated covariance (16) and model costs. To effectively relax this constraint, an automatic model selection (AMS) approach was introduced in [9] to perform multiple optimizations over all subsets of models and select the subset with minimum variance. Models not in the optimal subset are assigned zero samples in the sample allocation formulation provided in Section II.B.

C. Multi-Model MC with Python (MXMCPy)

The existence of several variations of multi-model MC approaches and the difficulty of knowing which will perform best for a given problem motivated the development and release of the open-source NASA software[†], Multi-Model MC with Python (`MXMCPy`) [9]. `MXMCPy` was the first publicly available library to offer convenient access to many of the existing multi-model MC algorithms, allowing users to easily assess and compare the performance of each to determine the best choice for a particular application. The software provides an efficient implementation of the solution to the sample allocation optimization problem in Equation (13), given the estimated run times, C_i, and covariance matrix (16) of the available models, and a specified computational budget for the analysis. An outline of the `MXMCPy` workflow is provided in Figure 1. With the optimal sample allocation, $\{\mathbf{x}\}$, prescribed by `MXMCPy`, a user performs the required evaluations of each model and can then form a multi-model MC estimator (10), again with functionality provided by `MXMCPy`. The numerical results to follow will demonstrate the application of this software for trajectory simulation.

III. Application

Viewing trajectory simulation as an uncertainty propagation problem as described in Section II.A, this work demonstrates the application of multi-model MC for pre-flight predictions of the ADEPT Sounding Rocket One (SR-1) test flight. This section provides a brief overview of the ADEPT test flight as well as the high-fidelity POST2 model used to simulate it, including the relevant QoIs to be predicted in the results section. A description of the high and low-fidelity models leveraged within the multi-model MC approach is also provided.

A. Modeling ADEPT SR-1 with POST2

ADEPT is a deployable aeroshell that can be deployed after launch to decelerate larger payloads to the ground when entering a planetary atmosphere. The first test flight of this technology, SR-1, was performed in September 2018 to demonstrate its ability to open exo-atmospherically and to characterize its stability during atmospheric flight [18]. A pre-flight trajectory simulation was performed and then compared to flight data from SR-1 [3]. Here, standard MC analysis was used to provide statistical predictions of various QoIs related to SR-1's trajectory during separation, apogee, and EDL.

[†] https://github.com/nasa/mxmcpy

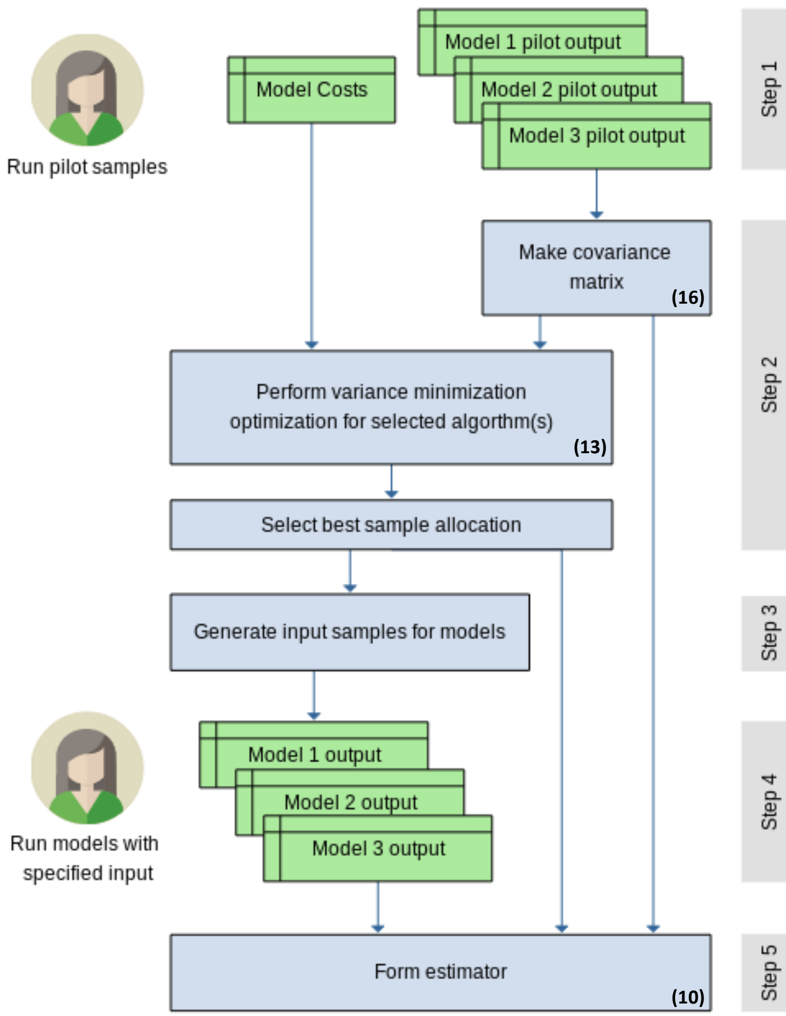


Fig. 1 `MXMCPy` workflow [9]. User responsibilities are indicated by green boxes while `MXMCPy` functionality is indicated by light blue boxes. Equations related to specific tasks are noted in the bottom right corner of boxes where appropriate.

Abbreviation	Description
lat-td	Geodetic latitude at touchdown
long-td	Longitude at touchdown
rllrt-80km	Roll rate at 80km geodetic altitude
rllrt-60km	Roll rate at 60km geodetic altitude
alt-apo	Apoapsis altitude
alt-mach1	Altitude at Mach 1
alt-deploy	ADEPT deploy altitude
time-td	Touchdown time
vel-term	Terminal velocity
accel-max	Max acceleration (Earth g's)
aoa-2km	Total angle of attack at 2km geodetic altitude
aoa-10km	Total angle of attack at 10km geodetic altitude
aoa-35km	Total angle of attack at 35km geodetic altitude
aoa-60km	Total angle of attack at 60km geodetic altitude
aoa-80km	Total angle of attack at 80km geodetic altitude

Table 1 Abbreviations and descriptions of the ADEPT QoIs considered in this work.

Model	High Fidelity	Reduced Physics	Coarse Time Step	Machine Learning
Run Time (sec)	2.19×10^{2}	4.74×10^{1}	2.80×10^{0}	7.00×10^{-4}

Table 2 Run times for the models used in this study.

In the context of the uncertainty propagation problem in Section II.A, the high-fidelity model, $\mathcal{M}_0$ in Equation (1), to simulate a single trajectory for SR-1 is the NASA Langley developed software, POST2 [10]. POST2 is a six degree-of-freedom flight mechanics simulation tool that can simulate trajectories of up to 20 independent or connected rigid bodies. The software has been used extensively in the past for several EDL missions [19–21]. Importantly, the resolution of POST2 simulations are controlled by a user-defined time step (Δt), where $\Delta t = 10^{-3}$ seconds was used for the trajectory simulations performed for ADEPT. The random inputs, $\mathbf{X}$, to the POST2 simulation represent parameters such as those governing atmosphere, aerodynamics, and mass properties of which there are 75 total ($d = 75$), where the marginal probability distributions, $p(\mathbf{x})$, describing each parameter are a mix of independent uniform and Gaussian distributions. The reader is referred to [3] for more details of the trajectory simulation specifications.

In this work, the ADEPT SR-1 trajectory simulation is revisited using the multi-model MC methodology described in Section II.B and compared with results from a standard MC approach. Fifteen QoIs are considered ($\mathbf{Y} \in \mathbb{R}^{15}$), including the predicted latitude and longitude at touchdown and roll rates and total angles of attack at various points along the trajectory, see Table 1 for a complete listing. For the same computational budget, Equation (8) for standard MC and Equation (12) for multi-model MC, estimators for the expected value of the various QoIs are constructed and the resulting MSE for both approaches is compared. The collection of models used in the multi-model MC approach is described next.

B. High and Low-Fidelity Models

To facilitate the use of a multi-model MC approach for ADEPT SR-1, a collection of $M + 1$ models for predicting the trajectory simulation QoIs in Table (1) is required. One model designates the high-fidelity model (Equation (1)) while the remaining M models are low-fidelity models (Equation (9)), designed to provide faster, but potentially inaccurate, QoI predictions. For illustration, three different classes of low-fidelity models are leveraged in this work. The collection of models are introduced and described below while the run time for each model is provided in Table 2.

High Fidelity Model

The high-fidelity model, $\mathcal{M}_0$, considered replicates the POST2 simulation used previously to generate pre-flight ADEPT

SR-1 predictions [3]. A time step of $\Delta t = 10^{-3}$ seconds is used.

Reduced Physics Model
This model is identical to the high-fidelity model but uses a simplified atmospheric model. In particular, it uses the U.S. standard atmospheric conditions rather than Earth Global Reference Atmospheric Model (GRAM) [22] and neglects any calculations pertaining to winds, resulting in over $4X$ speedup.

Coarse Time Step Model
This model is identical to the high-fidelity model but uses a time step that is two orders of magnitude larger, $\Delta t = 10^{-1}$.

Machine Learning Model
This model uses support vector machine (SVM) regression to build a surrogate model that approximates POST2 trajectory simulation. Here, the high-fidelity model was first run 250 times with randomly sampled input parameters to generate a training dataset along with the resulting QoIs from those simulations. Fifteen individual SVM regressors were then fit to the data for each QoI using the Scikit-Learn Python module [23]. Note that the run time reported in Table 2 is the total time required to evaluate all individual regressors for the full set of QoIs.

The correlation of each low-fidelity model with respect to the high-fidelity model was estimated for each QoI from 10,000 random samples and is shown in Figure 2. Generally speaking, a decrease in low-fidelity model run time is accompanied by a corresponding decrease in correlation to the high-fidelity model (faster models are less correlated). However, there is significant variation in the level of correlation across the QoIs, indicating that the performance of multi-model MC will vary for predicting different QoIs as well. In particular, all three low-fidelity models are highly correlated (≥ 0.85) for QoIs including `lat-td`, `long-td`, and `rllrt-80km` while low correlation (≤ 0.5) is observed for `accel-max` and `aoa-2km`. It is noted that the three models considered here were chosen mainly to illustrate the versatility of the multi-model MC approach to leverage various classes of low-fidelity models and a more thorough study to develop models with decreased run time and increased correlation was not performed. This type of analysis could yield significantly increased performance and represents a worthwhile avenue for future work.

IV. Results

This section provides a quantitative comparison of multi-model MC and standard MC applied to the ADEPT SR-1 trajectory simulation. The low-fidelity models introduced in the previous section are used in conjunction with the `MXMCPy` software to generate multi-model MC estimators for the QoIs listed in Table 1. The overarching goal is to study the accuracy improvements provided by a multi-model MC approach relative to a standard MC approach for trajectory simulation as the allotted computational budget is varied. These comparisons are done both with respect to a MC approach that uses the high-fidelity POST2 model as well as an approach that relies on the low-fidelity machine learning model, a common practice in computationally intensive uncertainty propagation problems.

Unless otherwise specified, two assumptions are made for the results to follow. First, in order to better control for the effect of errors in the estimated model covariance matrix on the multi-model MC estimators, Equation (16) is approximated using a large collection of 10,000 pilot samples. Second, the 250 POST2 simulations that were used to construct the machine learning low-fidelity model are considered an *offline* cost of the approach. In other words, the run time required for those simulations is not factored into the total costs of the multi-model MC and machine learning model-based MC results. Both of these assumptions are subsequently revisited to study their impact.

A. QoI Variance Reduction Study

First, a detailed study was done to determine the best multi-model MC estimators (i.e., those with the minimum estimator variance) for the various trajectory simulation QoIs using `MXMCPy`. The computational budget was fixed at $C^* = 10^4$ seconds for illustration and the sample allocation optimization problem (13) was solved for each available algorithm in `MXMCPy`. This was performed for two cases: 1) *Vector Opt.*: the vector QoI formulation provided by Equation (17) was optimized to find the single optimal sample allocation that minimized the variance across all QoIs simultaneously and 2) *Scalar Opt.*: the sample allocation and minimum estimator variance was optimized for each QoI

QoI	Vector Opt.	Scalar Opt. (Best)	Models Used - Scalar Opt.
lat-td	10.17	20.23	HiFi, CoarseDt, ML
long-td	9.09	16.97	HiFi, CoarseDt, ML
rllrt-80km	53.44	286.54	HiFi, RedPhys, ML
rllrt-60km	62.06	92.03	HiFi, ML
alt-apo	13.35	322.12	HiFi, CoarseDt, ML
alt-mach1	5.78	8.62	HiFi, CoarseDt, ML
alt-deploy	4.16	5.52	HiFi, CoarseDt, ML
time-td	2.52	3.23	HiFi, CoarseDt, ML
vel-term	1.15	1.43	HiFi, CoarseDt, ML
accel-max	0.73	1.03	HiFi, ML
aoa-2km	1.49	1.81	HiFi, CoarseDt, ML
aoa-10km	1.27	1.57	HiFi, CoarseDt, ML
aoa-35km	0.84	1.11	HiFi, CoarseDt, ML
aoa-60km	0.76	1.03	HiFi, CoarseDt, ML
aoa-80km	0.87	1.12	HiFi, CoarseDt, ML

Table 3 Variance reduction for multi-model MC relative to standard MC along with the optimal subset of models used to provide the best estimator for the *Scalar Opt.* case.

individually. For the *Vector Opt.* case, the scaling factors in Equation (17) were set as follows

$$\gamma_i = \frac{1}{(\max[Y_i] - \min[Y_i])^2}, \tag{18}$$

so that each QoI would be given equal importance for the sample allocation optimization. For the *Scalar Opt.* case, the AMS process was also performed with `MXMCPy` to determine the optimal subset of models for constructing estimators. The optimal estimator variance for multi-model MC using Equation (15).

The resulting variance reduction provided by multi-model MC (Equation (15)) versus standard MC (Equation (7)) from this study across all QoIs is shown in Figure 3.[‡] Since the *Vector Opt.* case optimizes all QoIs simultaneously in an average sense while the *Scalar Opt.* case represents a best-case estimator that targets each QoI exclusively, these results show the range of variance reduction possible using `MXMCPy` for the four models considered. The variance reduction achieved varies substantially across the QoIs and illustrates a direct correspondence to the degree of correlation observed between the models in Figure 2. Roughly two orders of magnitude or greater improvement is observed for QoIs such as `rllrt-80km` that exhibit high inter-model correlation. On the other hand, the variance reduction is ≈ 1.0 (performance comparable to standard MC) for `accel-max` and several of the QoIs representing total angle of attack. The delineation observed in performance across the QoIs considered is roughly between static quantities like landing location latitude/longitude that show high correlation and variance reduction versus more dynamic quantities like angle of attack late in the trajectory that show low correlation and variance reduction. This finding suggests that further low fidelity model development is required for more accurate approximations of these dynamic QoIs.

The variance reduction for multi-model MC versus standard MC is shown in more detail in Table 3 for each QoI. The rightmost column also indicates the subset of models that were selected from the AMS procedure that resulted in the best estimator for the *Scalar Opt.* case. It is interesting to note that in all cases there is at least one model that is left out and this leads to varying improvements to variance reduction. Additionally, in all but one case, the reduced physics model is omitted from the multi-model MC estimator, an indication that it is the least beneficial model overall. It can be seen that the worst performing QoIs in the *Scalar Opt.* case (`accel-max` and `aoa-2km`) offer only a 3% variance reduction versus standard MC.

Figure 4 shows the percentage of the computational budget allocated to each of the four models for the multi-model MC estimators, qualitatively illustrating the sample allocation prescribed by `MXMCPy` from the solution of Equation (13).

[‡]The optimal performing `MXMCPy` algorithms for all QoIs were the generalized ACV estimators proposed in [8].

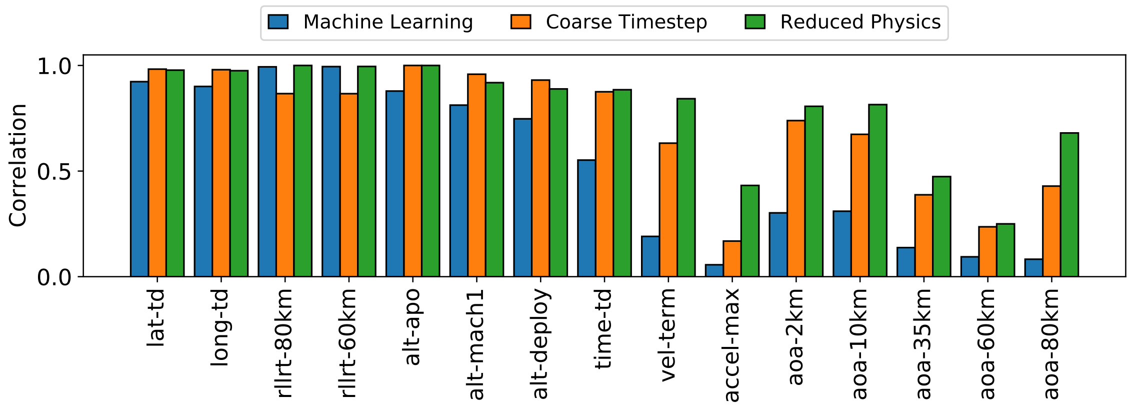


Fig. 2 Estimated correlations of low fidelity models versus the high fidelity model.

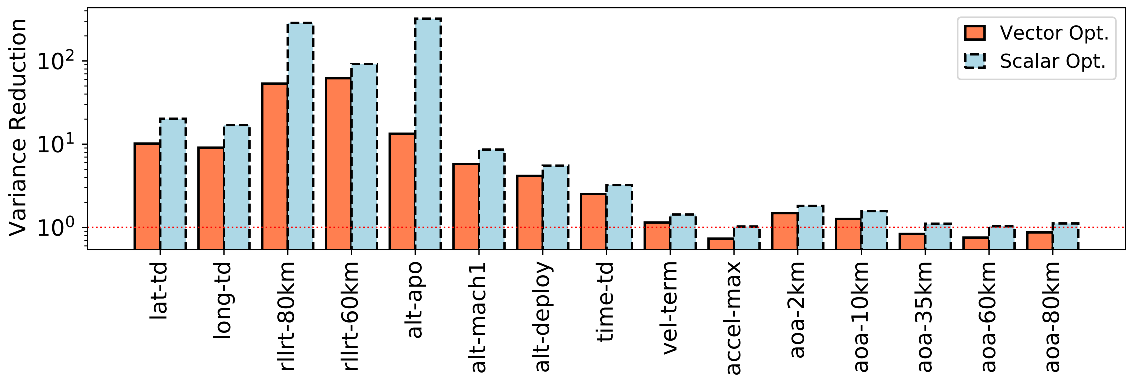


Fig. 3 Variance reduction relative to standard MC obtained by multi-model MC for each QoI. The vector optimization approach that minimizes estimator variance across all QoIs simultaneously is compared with the scalar optimization approach that minimizes the estimator variance of each QoI individually. The dotted red line at a variance reduction of 1.0 indicates equivalent performance between multi-model and standard MC estimators.

The left-most bar on the graph indicates the run time allocation resulting from the *Vector Opt.* case while the remaining bars represent the case of optimizing for each QoI individually in the *Scalar Opt.* case. The optimal resource allocation varies significantly across the QoIs, mirroring the relative amount of variation in inter-model correlation observed in Figure 2. It can again be seen that the reduced physics model is the least utilized model across the different QoIs. Interestingly, `MXMCPy` finds that the optimal sample allocation for the `accel-max` QoI is nearly identical to standard MC (all high-fidelity model evaluations), likely a result of the low correlation of the low-fidelity models observed in Figure 2. Note that while the machine learning model represent a relatively small amount of run time %, the number of model evaluations is generally orders of magnitude larger than the other models due to its computational efficiency.

B. QoI Estimator Accuracy

In order to evaluate the accuracy of multi-model MC estimators versus standard MC estimators, the MSE was estimated for all QoIs for a range of computational budgets. The reference solution used for comparison, denoted by $\bar{Y}_{ref}$, was calculated using a standard MC estimator with $N_0 = 10^5$ samples of the high-fidelity model, or a computational

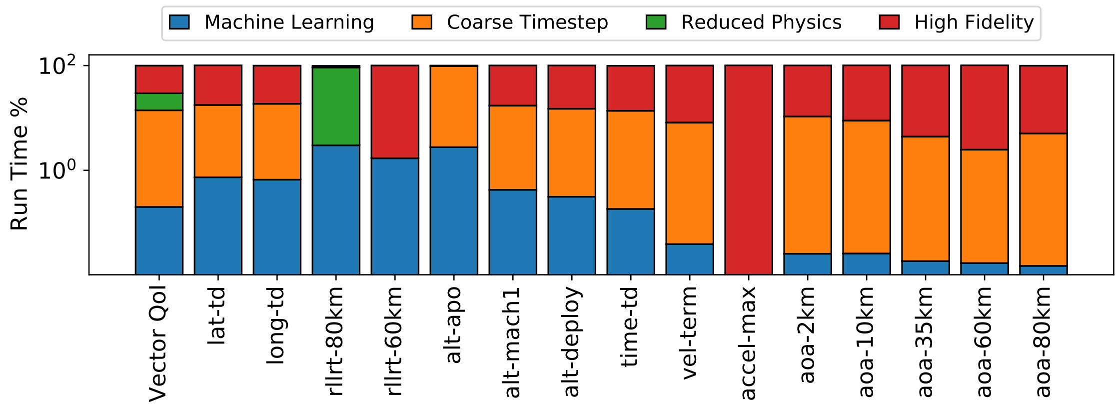


Fig. 4 Multi-model MC resource allocation by QoI - the percentage of computational budget spent running each model to generate estimators.

budget of $C^* = 2.19 \times 10^7$ seconds. The MSE of an estimator, $\bar{Y}$, was then estimated over ten random trials as

$$e(\bar{Y}) \approx \frac{1}{10} \sum_{i=1}^{10} \left(\bar{Y}_{ref} - \bar{Y}^{(i)} \right)^2, \tag{19}$$

where each of the estimates, $\bar{Y}^{(1)}, ..., \bar{Y}^{(10)}$, were computed using independent samples. Equation (19) was calculated for each of the QoIs in Table 1 for multi-model MC estimators using Equation (10) as well as standard MC estimators using Equation (3) with both the high-fidelity model (*high fidelity MC*) and machine learning-based low-fidelity model (*low fidelity MC*). The sample allocation for the *Vector Opt.* case was used for the multi-model MC estimators, representing a baseline performance possible with `MXMCPy`. Note that in this case all three low-fidelity models were leveraged (the AMS capability was not used).

The MSE calculated for each estimator versus total run time is shown for each QoI in Figure 5. Note that the multi-model MC estimators are required to evaluate the high-fidelity model at least once to ensure they are unbiased, so the MSE curves for both these estimators as well as the high-fidelity MC estimators are provided for higher run times only. A general trend observed for the low fidelity MC approach is that these QoI estimators provide the most accuracy for low computational budgets before being surpassed by the other estimators as the run time increases. This is because they are able to reduce estimator variance (Equation (6)) quickly since large number of model evaluations can be performed quickly, but then plateaus since there is an irreducible estimator bias (Equation (5)) present based on discrepancies with respect to the high-fidelity model predictions. On the other hand, the unbiased high fidelity MC estimators continue to converge towards the reference solution as the run time is increased, albeit with a significantly larger offset towards longer run times.

The MSE curves for the multi-model MC estimators illustrate the ability to provide efficient, unbiased predictions of the trajectory simulation QoIs. Whereas the low-fidelity MC estimators suffer from bias errors, the multi-model MC estimators are more accurate for moderate run times and converge to the reference solutions as run time is increased. With respect to high-fidelity MC, an order of magnitude or more improvement in MSE is observed for several of the QoIs (`lat-td`, `long-td`, `alt-apo`, `rllrt-80km`) for the same computation time. While there is significant variation in performance across the QoI estimators, the multi-model MC estimators are comparable to the accuracy of standard MC in the worst cases (`accel-max`, `aoa-2km`, `aoa-10km`, `aoa-2km`).

Recalling that multi-model MC leverages correlation among models to reduce estimator variance, the observed variability in performance here is a natural consequence of the high-to-low fidelity model correlation behavior observed in Figure 2. Comparing this figure versus the MSE results in Figure 5 illustrates a clear connection between the high inter-model correlation and significantly improved QoI estimator MSE and visa versa. Therefore, an important topic of future work is the development of low-fidelity models that exhibit higher correlation with the high-fidelity model for the lower performing QoIs. Furthermore, it is reiterated that superior performance is possible if one is interested in predicting only a single QoI (or fewer QoIs), since the optimal sample allocation would specifically target reducing variance in that QoI when solving Equation (13).

In order to further investigate the behavior of multi-model MC estimators and to revisit the assumptions made in the previous results, estimators for ADEPT SR-1 touchdown location variables (`lat-td`, `long-td`) were generated for computational budgets of 10^3, 10^4, and 10^5 seconds. Figure 6 compares the optimal estimator variance for each QoI resulting from the solution of Equation (13) with the resulting estimated MSE from Equation (19). Here, good agreement is observed between the predicted variance and the estimated MSE, verifying that the multi-model MC estimators are unbiased.

Figures 6 (b) and (d) also illustrate the effect of the number of pilot samples used to estimate the model covariance matrix (Equation (16) on the resulting estimator accuracy. Relatively little variation in the MSE behavior is observed when using 10,000 pilot samples, as was done in the previous results, versus fewer pilot samples (1000, 100, 10). Overall, the results for all numbers of pilot samples tested showed around $10X$ decrease in MSE for the same run time versus standard MC. Alternatively, Figure 6 shows that multi-model MC can yield the same accuracy in predicting landing location as standard MC with over an order of magnitude computational speedup.

Finally, Figure 7 shows the MSE versus run time for `lat-td` and `long-td` estimators when taking into account the run time required to train the machine learning low-fidelity models. Here, two MXMC cases are shown: one that uses all four models introduced in Section III.B and another that excludes the machine learning low-fidelity model for a total of three models. Recall that 250 high-fidelity model evaluations were performed to generate the SVM machine learning model and this was not factored into the previous results. Taking this cost into account, the relative performance gains from using a low-fidelity MC approach as well as a multi-model MC approach that includes the machine learning model are diminished. However, the multi-model MC approach still quickly surpasses the accuracy of high-fidelity MC. Additionally, it can be seen that using multi-model MC without the machine learning model still yields significant improvement over standard MC and can produce predictions for lower run times without the need to evaluate the high-fidelity model up front for training data.

V. Conclusion

This work explored the use of multi-model MC for increasing the speed and accuracy of trajectory simulation for EDL applications. Compared with a typical MC simulation approach that relies exclusively on an expensive, high-fidelity EDL model *or* a cheaper low-fidelity model, the proposed approach combines predictions from multiple models of varying fidelity and cost for efficient and unbiased uncertainty propagation. The POST2 flight mechanics simulation software was used for high-fidelity trajectory simulation while three types of low-fidelity models were considered, all with varying computational cost and correlation with respect to the high fidelity model. The `MXMCPy` software was leveraged for prescribing the optimal resource allocation across the available models to maximize estimator accuracy and subsequently combine predictions from each model to produce a single estimator for trajectory simulation QoIs.

A pre-flight analysis of the ADEPT SR-1 test flight was revisited to quantify performance improvements offered by a multi-model MC approach relative to standard MC. Expected value estimators were calculated for several QoIs describing trajectory state at various times of flight while accounting for uncertainties in parameters such as those governing the atmosphere and aerodynamics. For the same amount of computation time, multi-model MC provided up to two orders of magnitude accuracy improvement versus standard MC using the high-fidelity model only for estimating several QoIs. Moreover, the unbiasedness of multi-model MC was highlighted, demonstrating improved performance versus using standard MC with a low-fidelity model only that suffers from limitations on accuracy due to estimator bias.

While multi-model MC showed the potential to deliver substantial accuracy improvements for several ADEPT SR-1 QoIs considered, there were some QoIs with accuracy only comparable to a standard high-fidelity MC approach. It was shown that the variation in performance observed across different QoIs was directly related to the amount of correlation that exists between QoI predictions of the low fidelity models and the high-fidelity model. Therefore, an important avenue for future work is to develop low fidelity models that display higher correlation for the QoIs that had lower accuracy improvement, e.g., max acceleration and total angle of attack. This model development could include more advanced data-driven machine learning algorithms as well as exploring different options for reducing fidelity within the POST2 software itself. Finally, this work focused exclusively on estimating the *expected value* of trajectory simulation QoIs to establish a proof of concept. In order to provide a more practical capability for EDL applications, however, another future research direction is to extend the multi-model MC approach to predict more advanced statistical quantities, including rare event probabilities, joint probability distributions, and confidence ellipses of landing location.

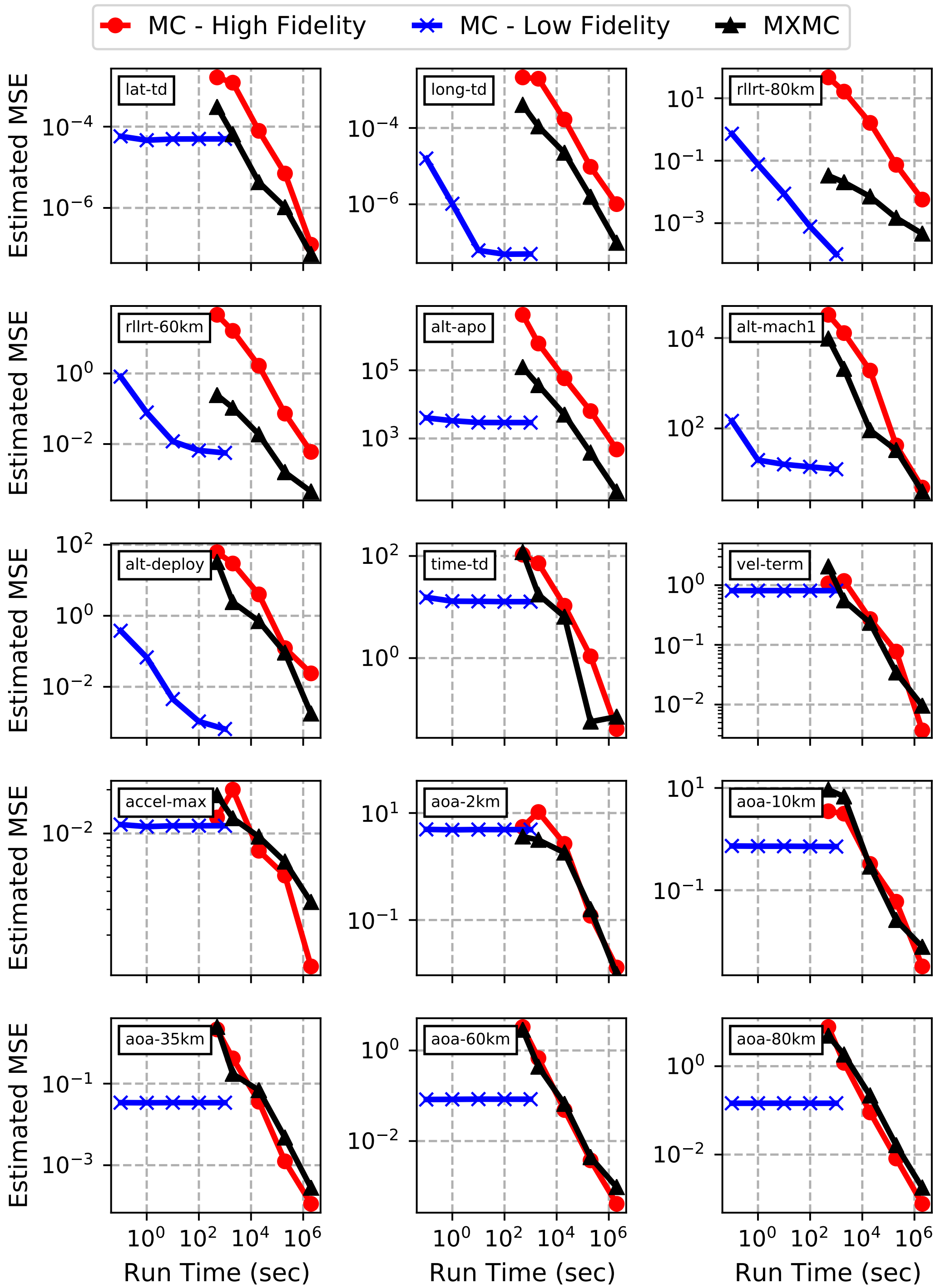
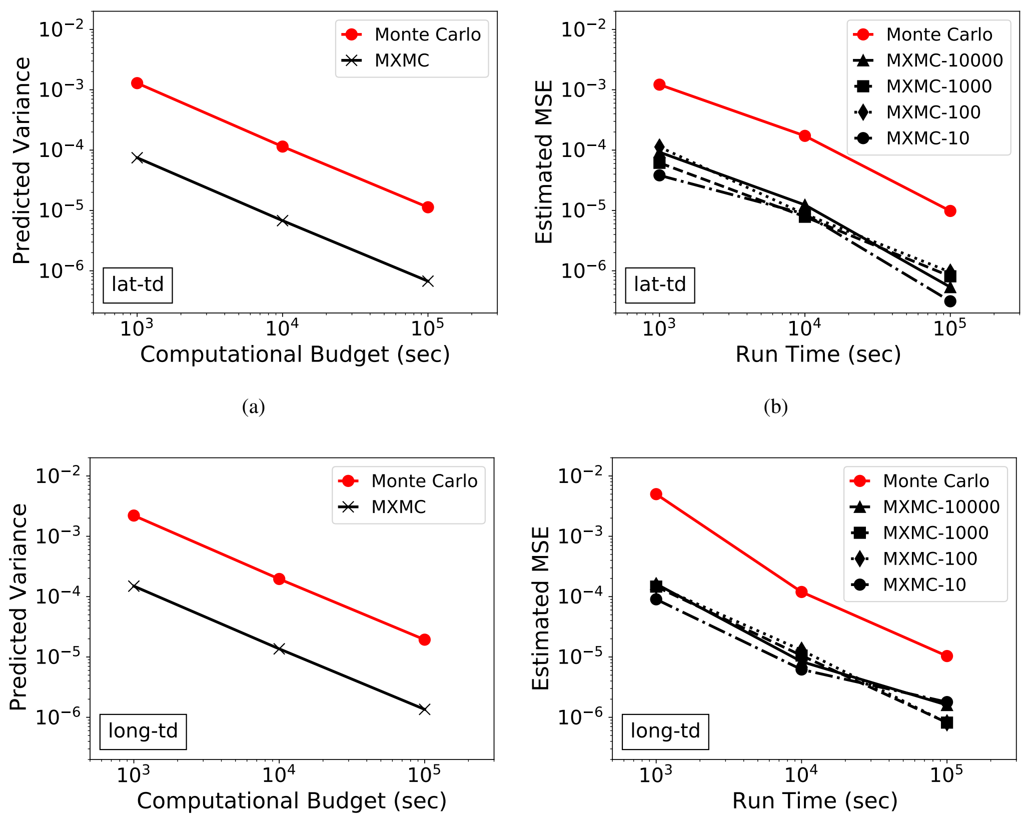


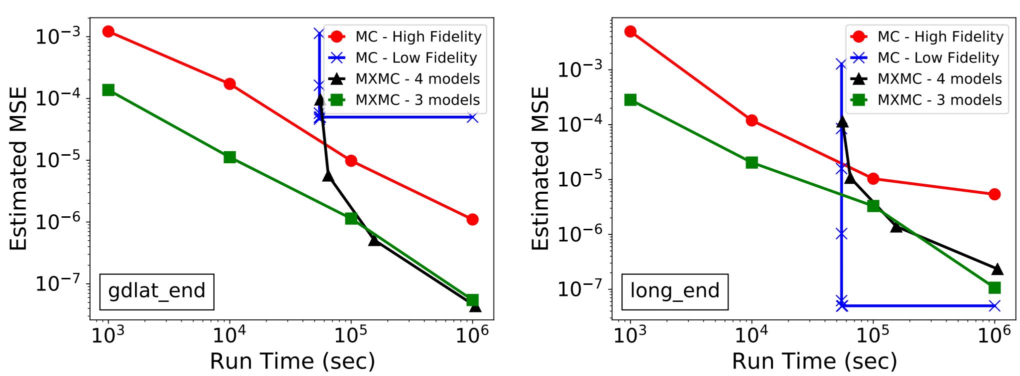
Fig. 5 MSE versus run time for all EDL QoIs comparing multi-model MC (MXMC) versus standard MC that uses the high-fidelity POST2 model and low-fidelity machine learning model.



(a) (b)

(c) (d)

Fig. 6 Predicted multi-model MC performance (variance from Equation (15) **on the left) versus actual performance (MSE from Equation** (19) **on the right) for ADEPT landing location. Figures (b) and (d) also illustrate the effect of number of pilot samples (10, 100, 1000, 10000) used to estimate the model covariance matrix.**



(a) (b)

Fig. 7 MSE versus run time for ADEPT landing location (`lat-td`, `long-td`) where the training time is factored into the surrogate model run time for both MC - Low Fidelity and MXMC - 4 models. Here, MXMC - 4 models results use all models introduced in Section III.B while MXMC - 3 models excludes the machine learning low-fidelity model.

VI. Acknowledgement

This work was supported by the NASA Internal Research and Development (IRAD), Center Innovation Fund (CIF), and NASA Internships and Fellowships (NIFS) Programs.

References

[1] Toups, L., and Hoffman, S., "Pioneering objectives and activities on the surface of Mars," *AIAA Space 2015 Conference*, AIAA, Pasadena, CA, 2015.

[2] Cianciolo, A. D., and Powell, R. W., "Entry, Descent, and Landing Guidance and Control Approaches to Satisfy Mars Human Mission Landing Criteria," *27th AAS/AIAA Space Flight Mechanics Meeting*, AIAA, San Antonio, TX, 2017.

[3] Dutta, S., and Green, J. S., "Flight Mechanics Modeling and Post-Flight Analysis of ADEPT SR-1," *AIAA Aviation 2019 Forum*, AIAA, Dallas, TX, 2019. doi:10.2514/6.2019-2900.

[4] Peherstorfer, B., Willcox, K., and Gunzburger, M., "Survey of multifidelity methods in uncertainty propagation, inference, and optimization," *SIAM Review*, Vol. 60, No. 3, 2018, pp. 550–591.

[5] Giles, M. B., "Multi-level Monte Carlo path simulation," *Operations Research*, Vol. 56, No. 3, 2008, pp. 607–617.

[6] Peherstorfer, B., Willcox, K., and Gunzburger, M., "Optimal Model Management for Multifidelity Monte Carlo Estimation," *SIAM Journal on Scientific Computing*, Vol. 38, 2016, pp. A3163–A3194. doi:10.1137/15M1046472.

[7] Gorodetsky, A., Geraci, G., Eldred, M., and Jakeman, J. D., "A generalized approximate control variate framework for multifidelity uncertainty quantification," *Journal of Computational Physics*, 2020, p. 109257. doi:https://doi.org/10.1016/j.jcp.2020.109257, URL `http://www.sciencedirect.com/science/article/pii/S0021999120300310`.

[8] Bomarito, G. F., Leser, P. E., Warner, J. E., and Leser, W. P., "On the Optimization of Approximate Control Variates with Parametrically Defined Estimators," *In Preparation*, 2020.

[9] Bomarito, G. F., Warner, J. E., Leser, P. E., Leser, W. P., and Morrill, L., "Multi Model Monte Carlo with Python (MXMCPy)," *NASA/TM-2020-220585*, 2020.

[10] Lugo, R., Shidner, J., Powell, R., Marsh, S., Hoffman, J., Litton, D. K., and Schmitt, T., "Launch vehicle ascent trajectory simulation using the Program to Optimize Simulated Trajectories II (POST2)," *Astronautical Sciences Spaceflight Mechanics 2017 Conference*, AIAA, San Antonio, TX, 2019. doi:AAS17-274.

[11] Rubinstein, R. Y., and Kroese, D. P., *Simulation and the Monte Carlo Method*, 3rd ed., Wiley Publishing, 2016.

[12] Razavi, S., Tolson, B. A., and Burn, D. H., "Review of surrogate modeling in water resources," *Water Resources Research*, Vol. 48, No. 7, 2012. doi:https://doi.org/10.1029/2011WR011527, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011WR011527`.

[13] Hesterberg, T., "Control variates and importance sampling for efficient bootstrap simulations," *Statistics and Computing*, Vol. 6, 1996, pp. 147–157. doi:10.1007/BF00162526.

[14] Rubinstein, R. Y., and Marcus, R., "Efficiency of Multivariate Control Variates in Monte Carlo Simulation," *Operations Research*, Vol. 33, No. 3, 1985, pp. 661–677. URL `http://www.jstor.org/stable/170564`.

[15] Quaglino, A., Pezzuto, S., and Krause, R., "High-dimensional and higher-order multifidelity Monte Carlo estimators," *Journal of Computational Physics*, Vol. 388, 2019, pp. 300–315.

[16] Qian, E., Peherstorfer, B., O'Malley, D., Vesselinov, V. V., and Willcox, K., "Multifidelity Monte Carlo estimation of variance and sensitivity indices," *SIAM/ASA Journal on Uncertainty Quantification*, Vol. 6, No. 2, 2018, pp. 683–706.

[17] Bierig, C., and Chernov, A., "Approximation of probability density functions by the multilevel Monte Carlo maximum entropy method," *Journal of Computational Physics*, Vol. 314, 2016, pp. 661–681.

[18] Tynis, J. A., Karlgaard, C. D., Williams, J. D., and Dutta, S., "Reconstruction of the Adaptable Deployable Entry and Placement Technology Sounding Rocket One Flight Test," *AIAA Aviation 2019 Forum*, AIAA, Dallas, TX, 2019. doi:10.2514/6.2019-2899.

[19] Braun, R. D., Powell, R. W., Engelund, W. C., Gnoffo, P. A., Weilmuenster, K. J., and Mitcheltree, R. A., "Mars Pathfinder Six-Degree-of-Freedom Entry Analysis," *Journal of Spacecraft and Rockets*, Vol. 32, No. 6, 1995, pp. 993–1000. doi:10.2514/3.26720.

[20] Desai, P. N., Qualls, G. D., and Schoenenberger, M., “Reconstruction of the Genesis Entry,” *Journal of Spacecraft and Rockets*, Vol. 45, No. 1, 2008, pp. 33–38. doi:10.2514/1.30042.

[21] Desai, P. N., and Qualls, G. D., “Stardust Entry Reconstruction,” *Journal of Spacecraft and Rockets*, Vol. 47, No. 5, 2010, pp. 736–740. doi:10.2514/1.37679.

[22] Leslie, F. W., and Justus, C. G., “The NASA Marshall Space Flight Center Earth Global Reference Atmosphere Model-2010 Version,” *NASA/TM-2011-216467*, 2011.

[23] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E., “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, Vol. 12, 2011, pp. 2825–2830.