SAND2022-4140C

Automating Model Selection and Tuning for Multifidelity UQ (MFUQ)

James Warner[1], Geoffrey Bomarito[1], Gianluca Geraci[2], Michael Eldred[2], Marten Thompson[1,3], John Jakeman[2], Patrick Leser[1], Paul Leser[1], Alex Gorodetsky[4]

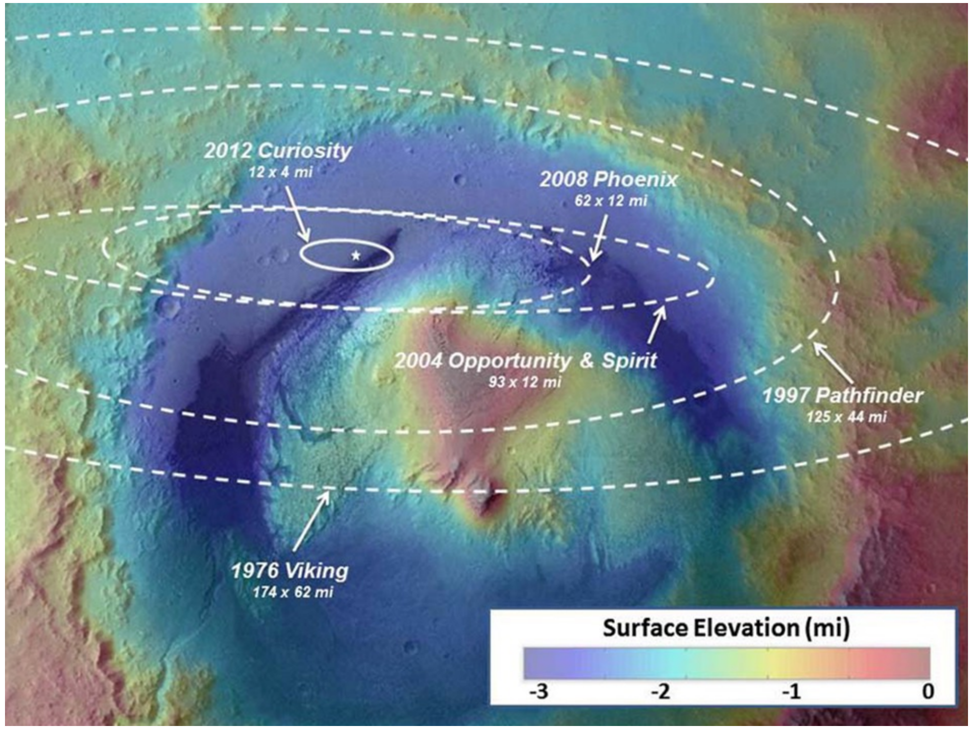
1. NASA Langley Research Center, Hampton, VA
2. Sandia National Laboratories, Albuquerque, NM
3. University of Minnesota Twin Cities, Minneapolis, MN
4. University of Michigan, Ann Arbor, MI

MS77: Advanced Multilevel and Multifidelity UQ Strategies: Applications, Generalized Model Hierarchies, and Data-Driven Approaches
SIAM Conference on Uncertainty Quantification
April 13th, 2022

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Motivation: Trajectory Guidance

- Orders of magnitude improvement in landing precision required to enable human Mars missions
  - **Long-term goal**: robust trajectory optimization under uncertainty

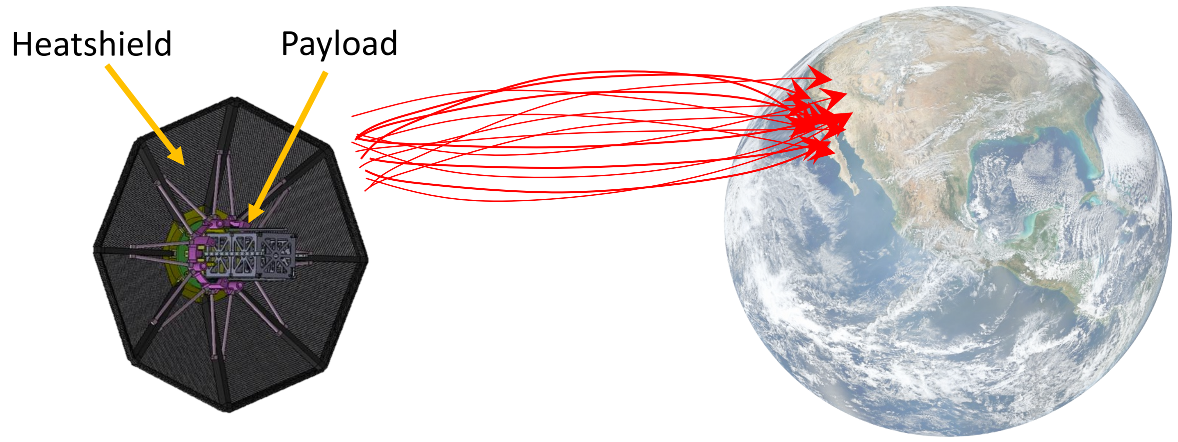


Mars landing precision over time

Motivation: Trajectory Guidance

- **Current focus:** uncertainty propagation for high-fidelity trajectory simulation
  - High dimensional inputs – e.g., atmosphere conditions, vehicle state
  - Multiple quantities of interest (QoIs) – e.g., landing location, flight time
  - Model: Program to Optimize Simulated Trajectories II (POST2)

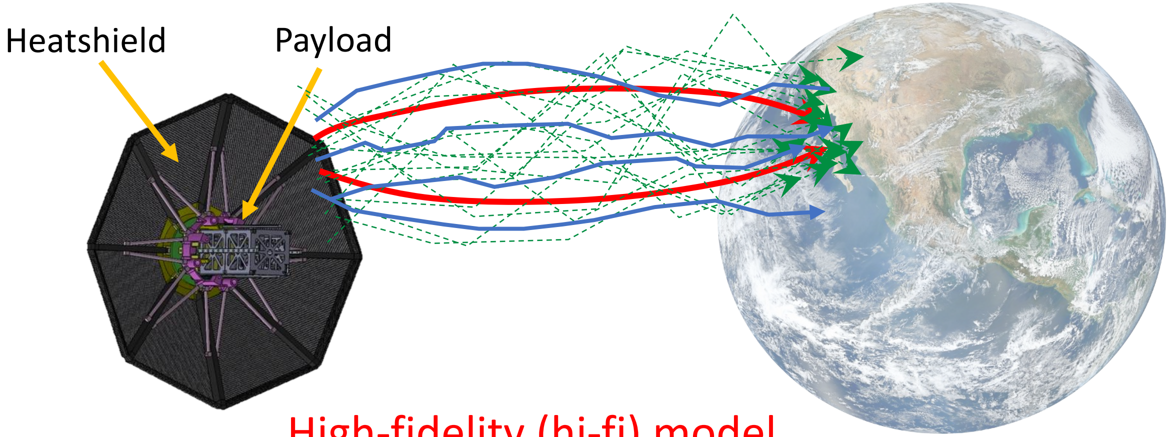
Monte Carlo (MC) simulation



Motivation: Trajectory Guidance

- **Current focus:** uncertainty propagation for high-fidelity trajectory simulation
  - High dimensional inputs – e.g., atmosphere conditions, vehicle state
  - Multiple quantities of interest (QoIs) – e.g., landing location, flight time
  - Model: Program to Optimize Simulated Trajectories II (POST2)

MFUQ

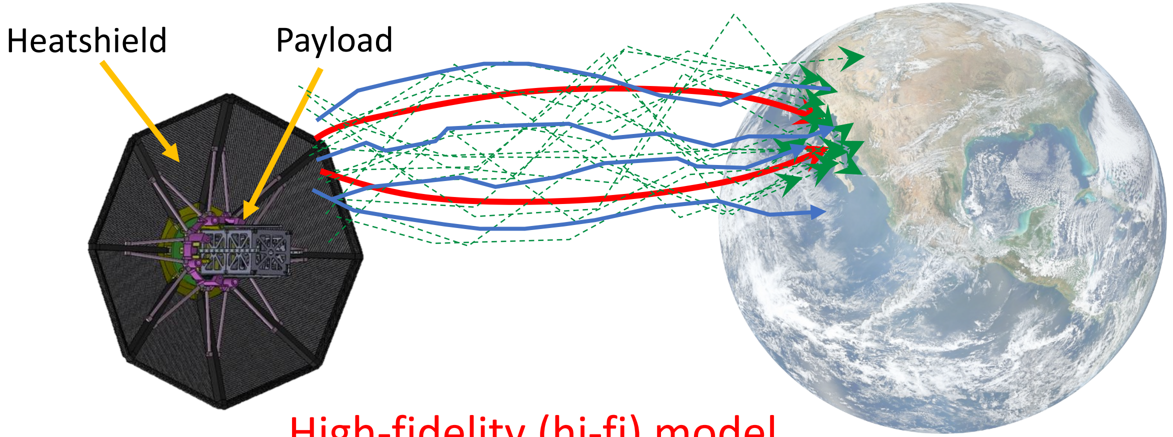


High-fidelity (hi-fi) model
Medium-fidelity (med-fi) model
Low-fidelity (lo-fi) model

Motivation: Trajectory Guidance

- **Current focus:** uncertainty propagation for high-fidelity trajectory simulation
  - High dimensional inputs – e.g., atmosphere conditions, vehicle state
  - Multiple quantities of interest (QoIs) – e.g., landing location, flight time
  - Model: Program to Optimize Simulated Trajectories II (POST2)

MFUQ

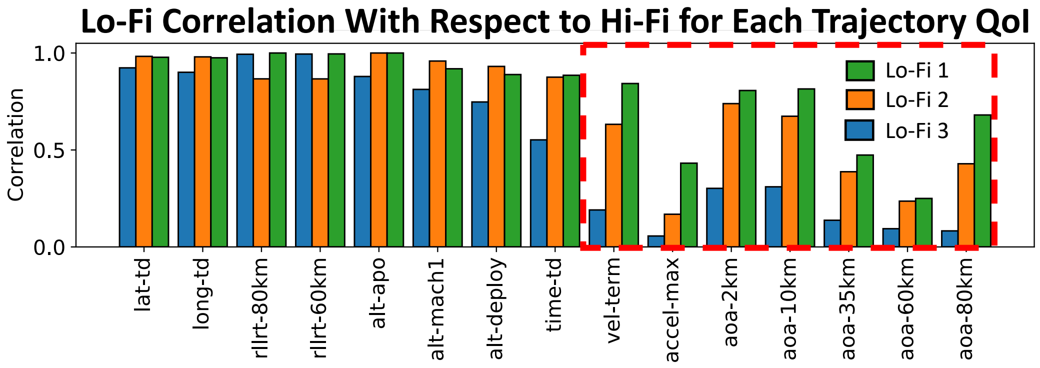


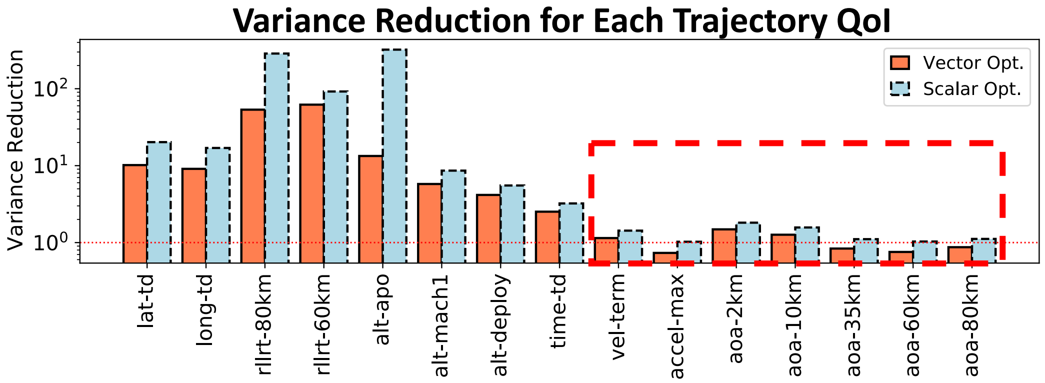
High-fidelity (hi-fi) model

Medium-fidelity (med-fi) model

Low-fidelity (lo-fi) model

AIAA SciTech 2021[1]:



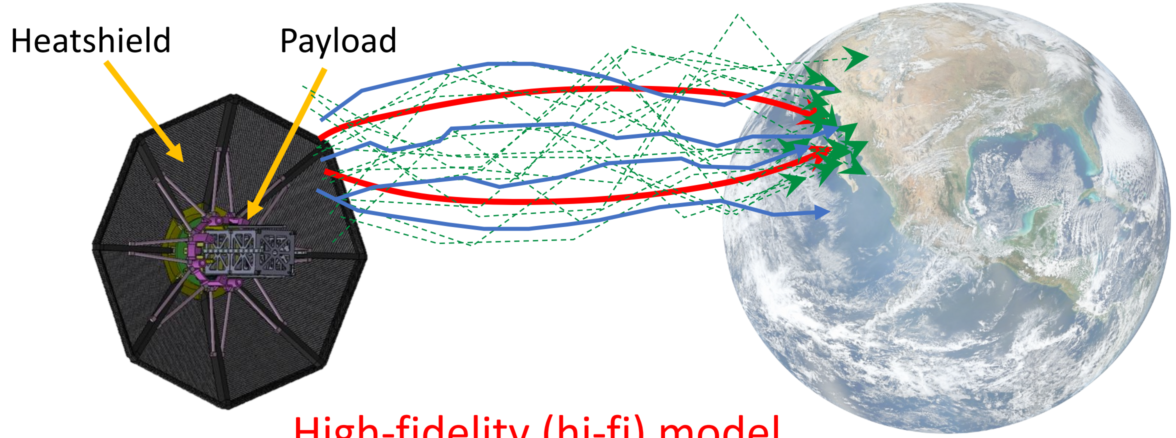


➢ **Sub-optimal lo-fi models = sub-optimal MFUQ**

[1] J. E. Warner et al. AIAA SciTech 2021.

Outline: Model Tuning for MFUQ

- Model tuning is important
- How we can do it optimally (In Progress)
- Application to trajectory simulation
  - Optimal time step selection for lo-fi models



High-fidelity (hi-fi) model
Medium-fidelity (med-fi) model
Low-fidelity (lo-fi) model

Approximate Control Variates (ACV)[1]

$$\tilde{Q} = \hat{Q}(z) + \sum_{i=1}^{M} \alpha_i \left(\hat{Q}_i(z_i^1) - \hat{Q}_i(z_i^2) \right)$$

$Q:$ hi-fi model
Q_i: lo-fi models
$\hat{Q}$: MC estimator
α_i: CV weights

- Multilevel Monte Carlo (**MLMC**)[2] and Multifidelity Monte Carlo (**MFMC**)[3] are instances of this estimator
- New ACV estimators[1] based on independent sampling (**ACV-IS**), multifidelity sampling (**ACV-MF**) and their corresponding generalizations[4] (**GIS**, **GMF**)
- Estimator is unbiased (wrt $E[Q]$)
- ACV estimator variance: $Var[\tilde{Q}] = Var[\hat{Q}](1 - R_{ACV}^2)$

[1] Gorodetsky, A A., et al. Journal of Computational Physics (2020)
[2] Giles, M B. Operations Research (2008)
[3] Peherstorfer, B, et al. SIAM Journal on Scientific Computing (2016)
[4] Bomarito, G. F., et al. Journal of Computational Physics (2022)

ACV-MF Estimator Variance

$$Var[\tilde{Q}] = Var[\hat{Q}](1 - R^2_{ACV-MF})$$

cov of lo-fi models w.r.t. each other

cov of lo-fi models w.r.t. hi-fi

sampling ratios

$$R^2_{ACV-MF}(\vec{r}) = \frac{1}{Var[Q]}[diag[\boldsymbol{F}(\vec{r})] \circ \boldsymbol{c}]^T [\boldsymbol{C} \circ \boldsymbol{F}(\vec{r})]^{-1} [diag[\boldsymbol{F}(\vec{r})] \circ \boldsymbol{c}]$$

Matrix representing ACV-MF sampling strategy

Variance minimization for sample allocation:

$$\underset{N,\vec{r}}{\text{argmin}}\ Var[\tilde{Q}](N,\vec{r}) \quad \text{s.t.} \quad W^{total}(N,\vec{r}) \leq W^{target}$$

Estimator cost

Computational budget

ACV-MF Estimator Variance

Introduce model tuning parameters $\vec{\beta}$

cov of lofi models w.r.t. each other

cov of lofi models w.r.t. hifi

$$R^2_{ACVMF}(\vec{r},\vec{\beta}) = \frac{1}{Var[Q]}\left[diag[\boldsymbol{F}(\vec{r})] \circ \boldsymbol{c}(\vec{\beta})\right]^T \left[\boldsymbol{C}(\vec{\beta}) \circ \boldsymbol{F}(\vec{r})\right]^{-1}\left[diag[\boldsymbol{F}(\vec{r})] \circ \boldsymbol{c}(\vec{\beta})\right]$$

In general, these are not known a priori and must be estimated

- Can reformulate generally in terms of only model correlations
- Model costs are also a function of $\vec{\beta}$ (potentially known)

Optimization Approach*

$$\underset{N,\vec{r},\vec{\beta}}{\text{argmin}} \; Var[\tilde{Q}](N,\vec{r},\vec{\beta}) \quad \text{s.t.} \quad W^{total}(N,\vec{r},\vec{\beta}) \leq W^{target}$$

- Build global surrogate for $\boldsymbol{c}(\vec{\beta})$ from pilot samples
  - N_{tun}: number of values of tuning parameters to investigate
  - N_{pilot}: number of pilot samples at each set of tuning parameters
  - Local quadratic interpolant
- Assume known relationship of model costs
- Gradient based-optimization (SLSQP)

*work in progress

- Currently based on single iteration from pilot samples for all N_{tun}, N_{pilot}

Analytical Example[1]

- **Goal**: study the effect of hyperparameter, θ_1, for lo-fi model, Q_1, on estimator variance

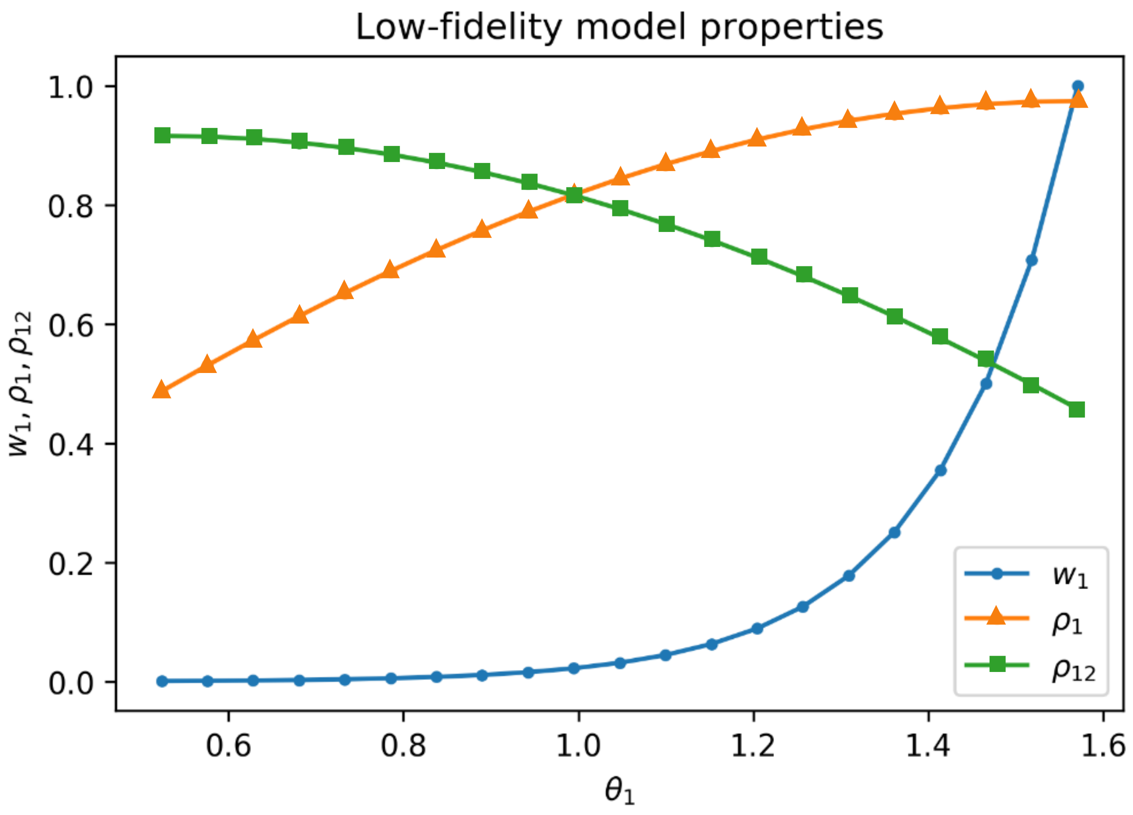
Hi-fi and lo-fi models

$$Q = A(\cos\theta x^5 + sin\theta y^5)$$
$$Q_1 = A_1(\cos\theta_1 x^3 + sin\theta_1 y^3)$$
$$Q_2 = A_2(\cos\theta_2 x + sin\theta_2 y)$$

Model costs:

$$w = 1$$
$$w_2 = 10^{-3}$$
$$\log w_1 = \log w_2 + \frac{\log w_2 - \log w}{\theta_2 - \theta}(\theta_1 - \theta_2)$$

A,A_1, A_2, θ, θ_2 are constants



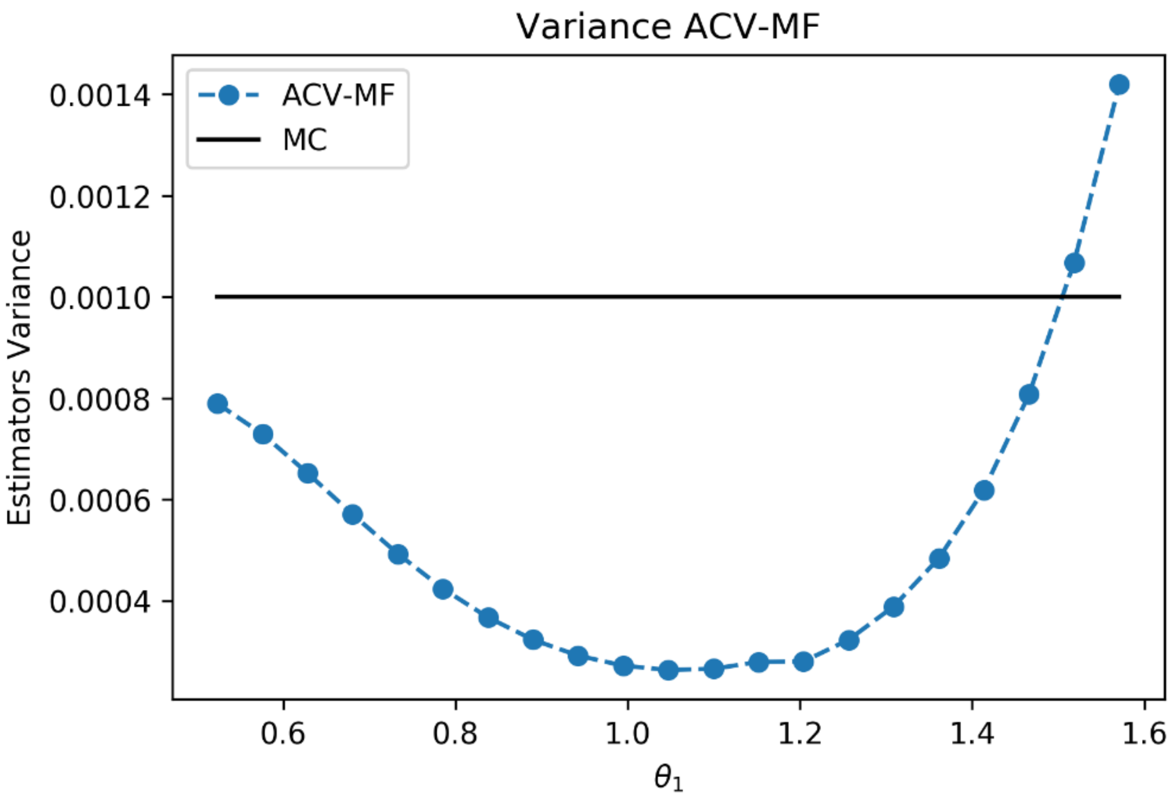
w_1: model cost

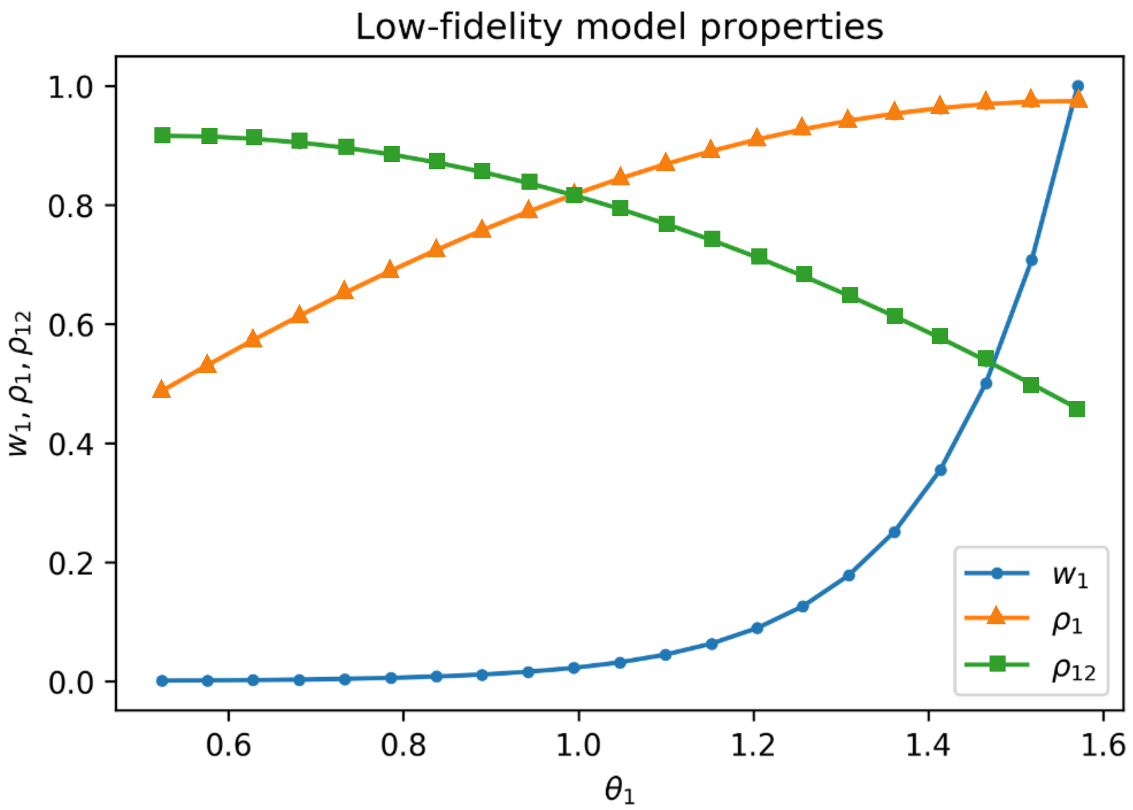
ρ_1 : correlation between Q_1 and Q

ρ_{12} : correlation between Q_1 and Q_2

[1] Gorodetsky, A A., et al. Journal of Computational Physics (2020)

Analytical Example[1]





w_1: model cost

ρ_1 : correlation between Q_1 and Q

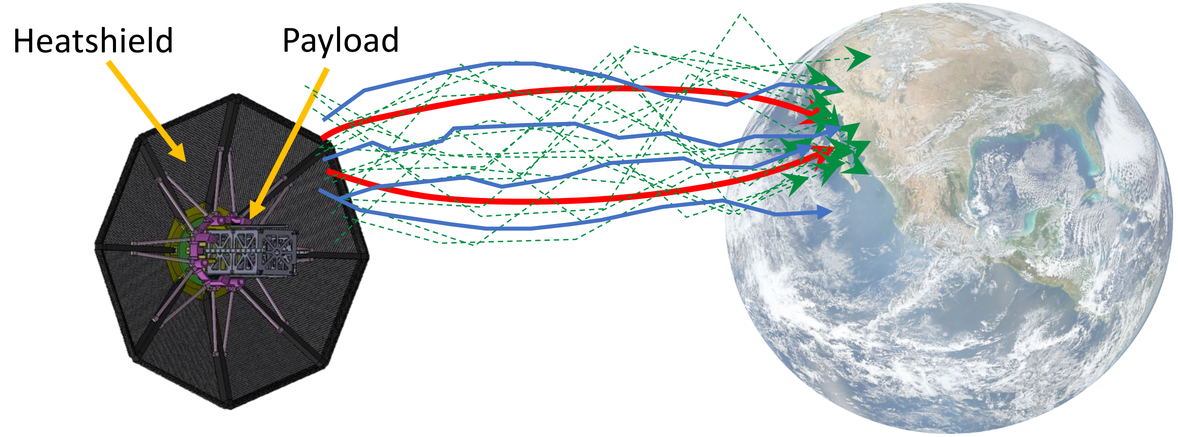
ρ_{12} : correlation between Q_1 and Q_2

Model tuning can greatly affect estimator variance

[1] Gorodetsky, A. A., et al. Journal of Computational Physics (2020)

Trajectory Simulation Example

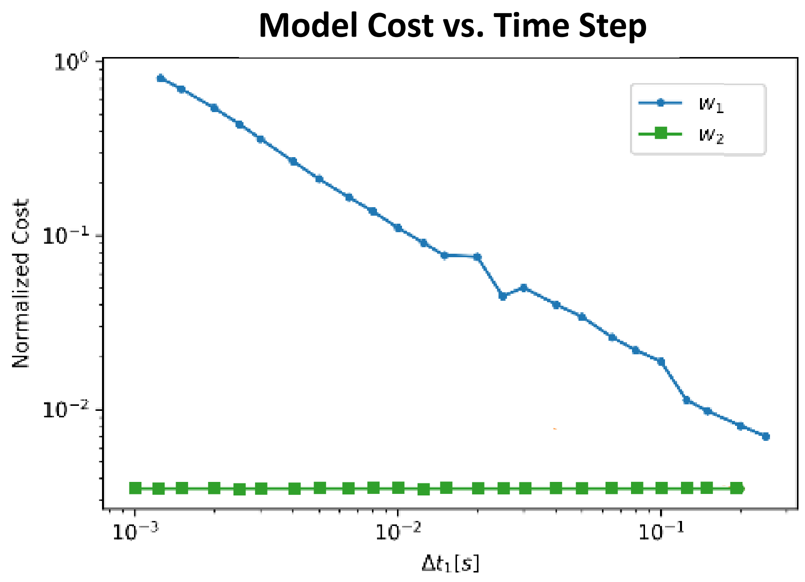
Goal: Predict the flight time of an umbrella heatshield reentering the Earth's atmosphere within computational budget

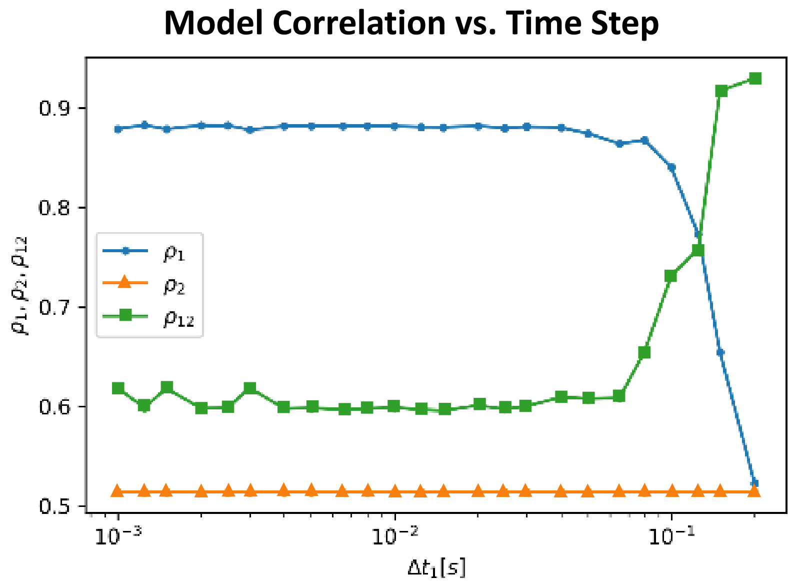


High-fidelity (Q): timestep = 0.001

Mid-fidelity (Q_1): timestep = $0.001 \leq \Delta t_1 \leq 0.25$

Low-fidelity (Q_2): timestep = 0.25





w_1: Q_1 model cost

ρ_1 : correlation between Q_1 and Q

ρ_{12} : correlation between Q_1 and Q_2

Trajectory Simulation Example

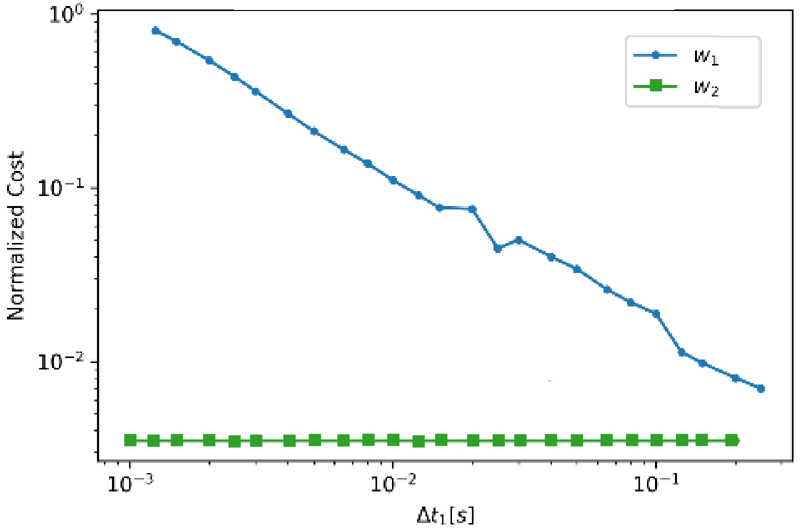
Goal: Predict the flight time of an umbrella heatshield reentering the Earth's atmosphere within computational budget

- Determine mid-fidelity time step through joint optimization:

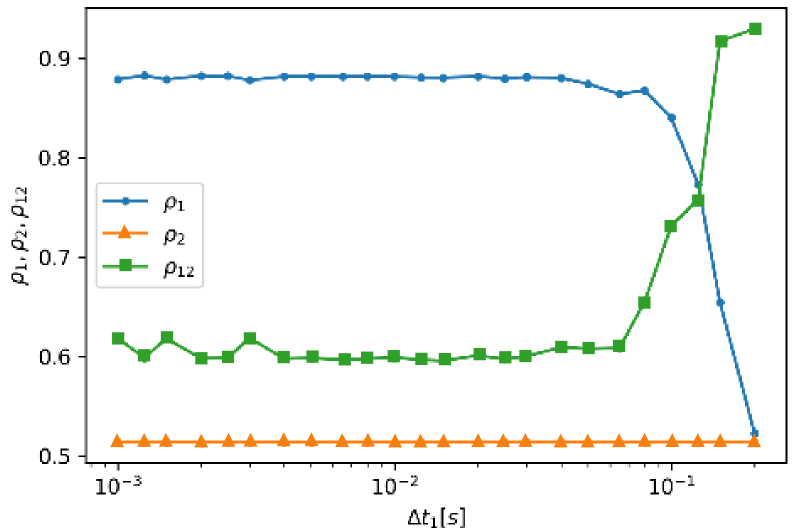
$$\underset{N, r_1, r_2, \Delta t_1}{\text{argmin}} \; Var[\tilde{Q}](N, r_1, r_2, \Delta t_1) \;\text{ s.t. }\; W^{total}(N, r_1, r_2, \Delta t_1) \leq W^{target}$$

- Compare correlation surrogate models with varying amounts of data
  1. $\mathrm{N_{tun}}$=24, $\mathrm{N_{pilot}} = 200$
  2. $\mathrm{N_{tun}}$=24, $\mathrm{N_{pilot}} = 50$
  3. $\mathrm{N_{tun}}$=6, $\mathrm{N_{pilot}} = 200$
  - Reference case (*Ref Surr*) with all data: $\mathrm{N_{tun}}$=24, $\mathrm{N_{pilot}} = 500$

Model Cost vs. Time Step



Model Correlation vs. Time Step

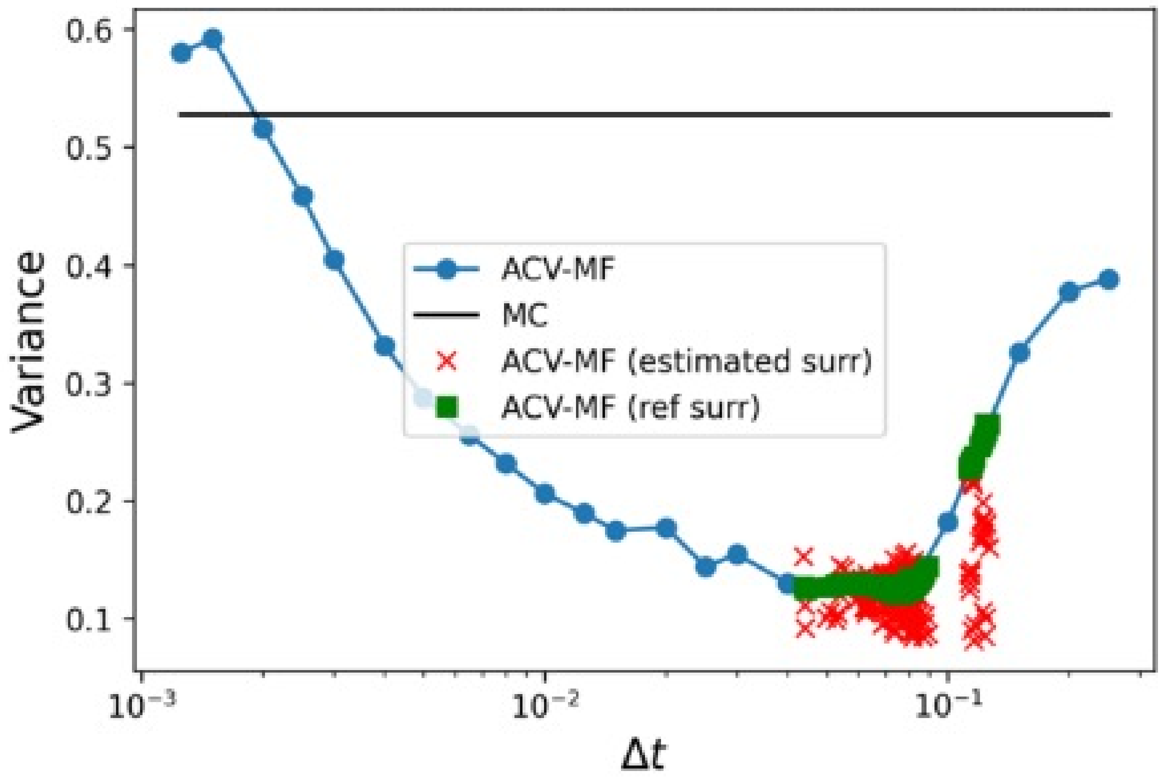


w_1: Q_1 model cost
ρ_1 : correlation between Q_1 and Q
ρ_{12} : correlation between Q_1 and Q_2

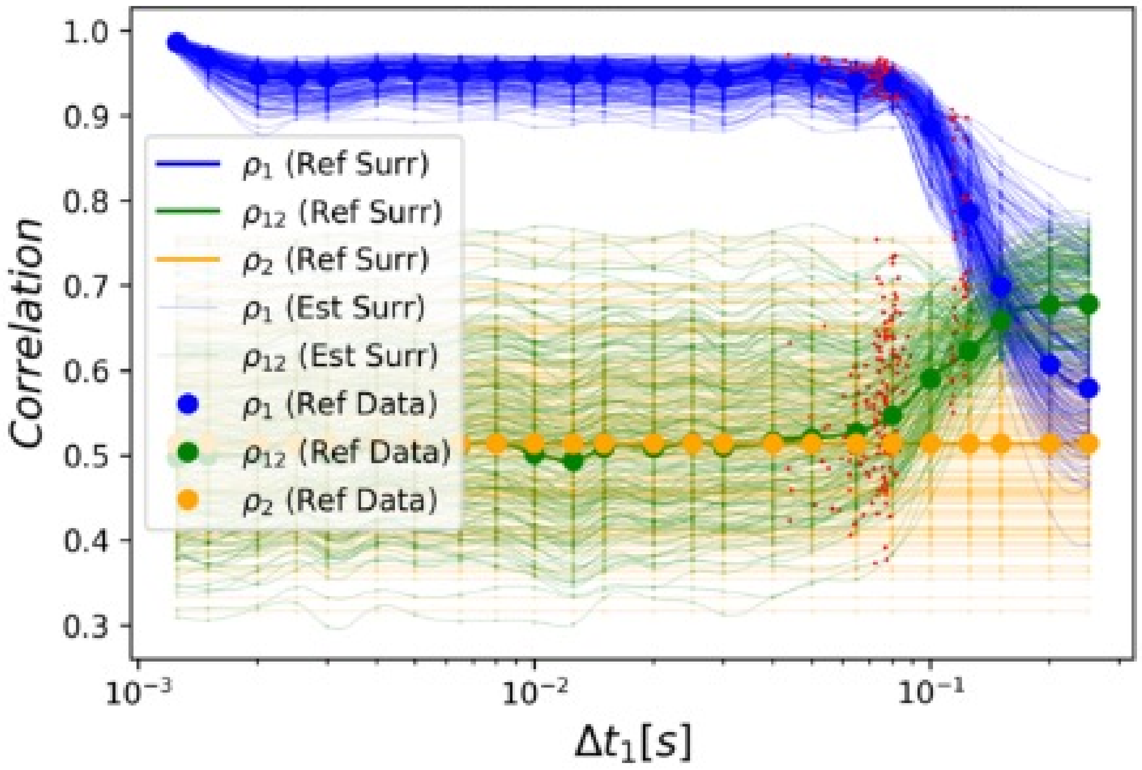
Results: Most Accurate Surrogate

- Optimization using $\mathrm{N_{tun}} = 24$, $\mathrm{N_{pilot}} = 200$ to build correlation surrogate models; 200 random trials

Estimator Variance vs. Time Step



Estimated Model Correlation vs. Time Step

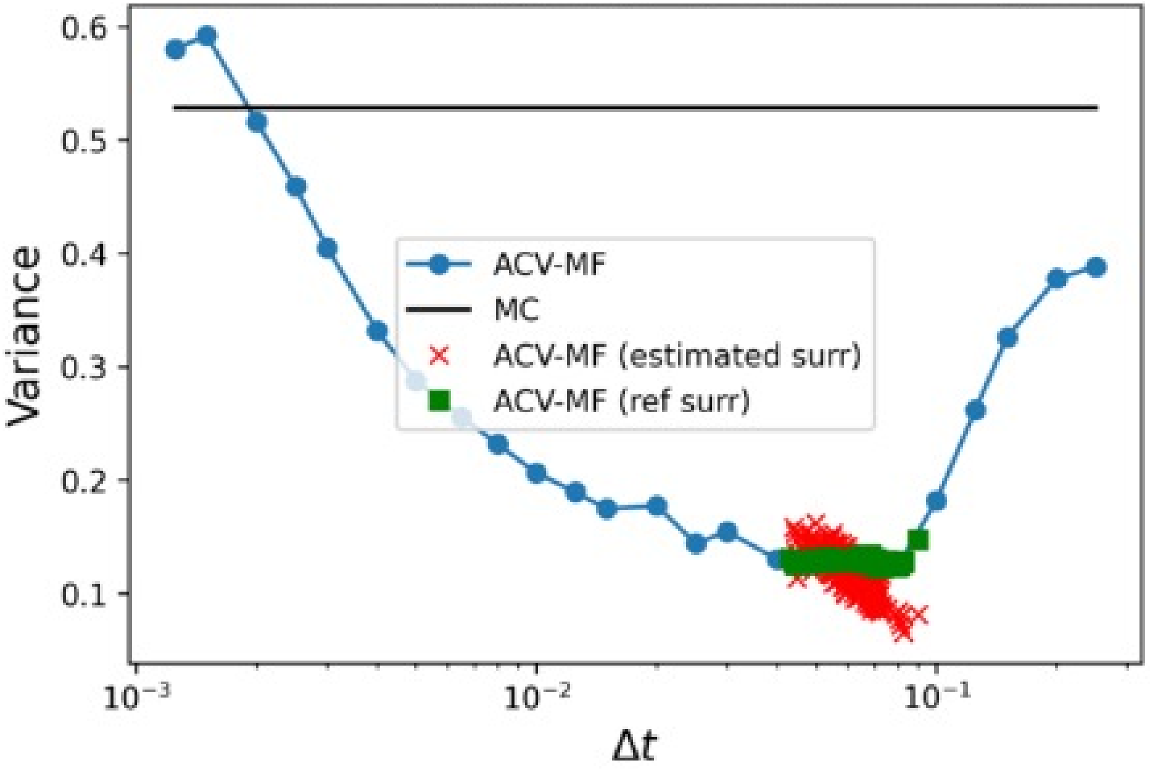


Optimal model tuning is achievable using surrogates for correlation

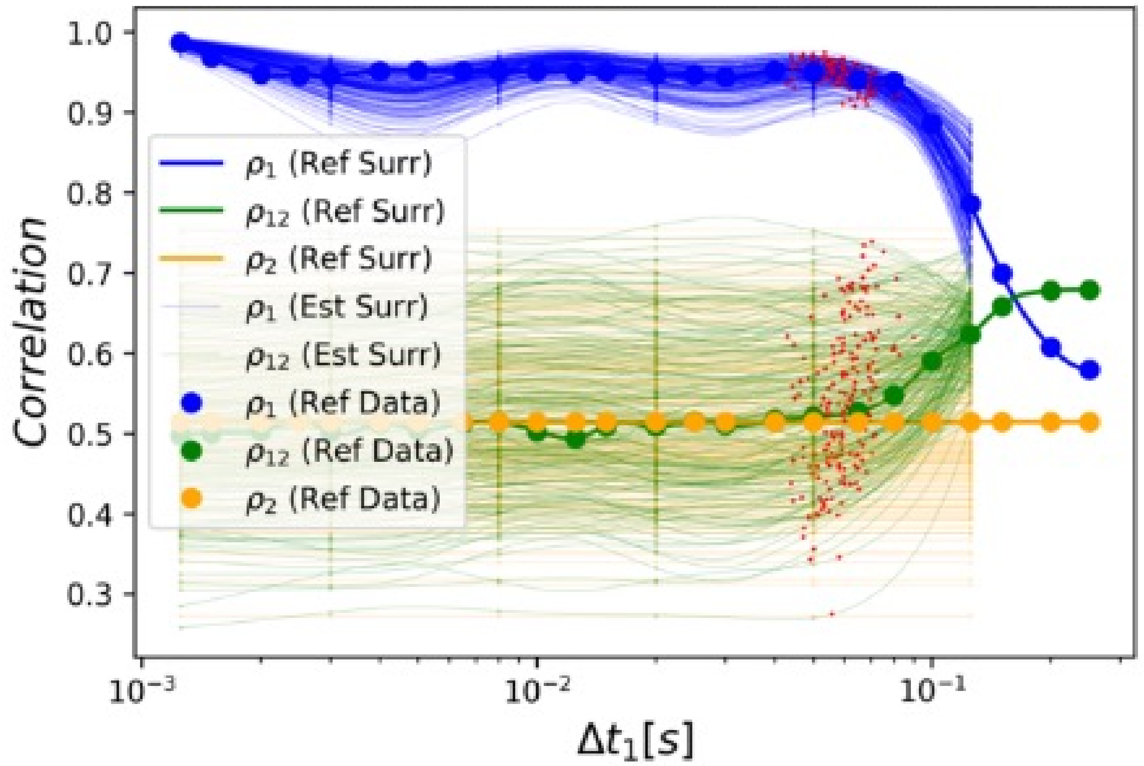
Results: Sparser Grid

- Optimization using $\mathbf{N_{tun} = 6}$, $\mathrm{N_{pilot}} = 200$ to build correlation surrogate models; 200 random trials

Estimator Variance vs. Time Step



Estimated Model Correlation vs. Time Step

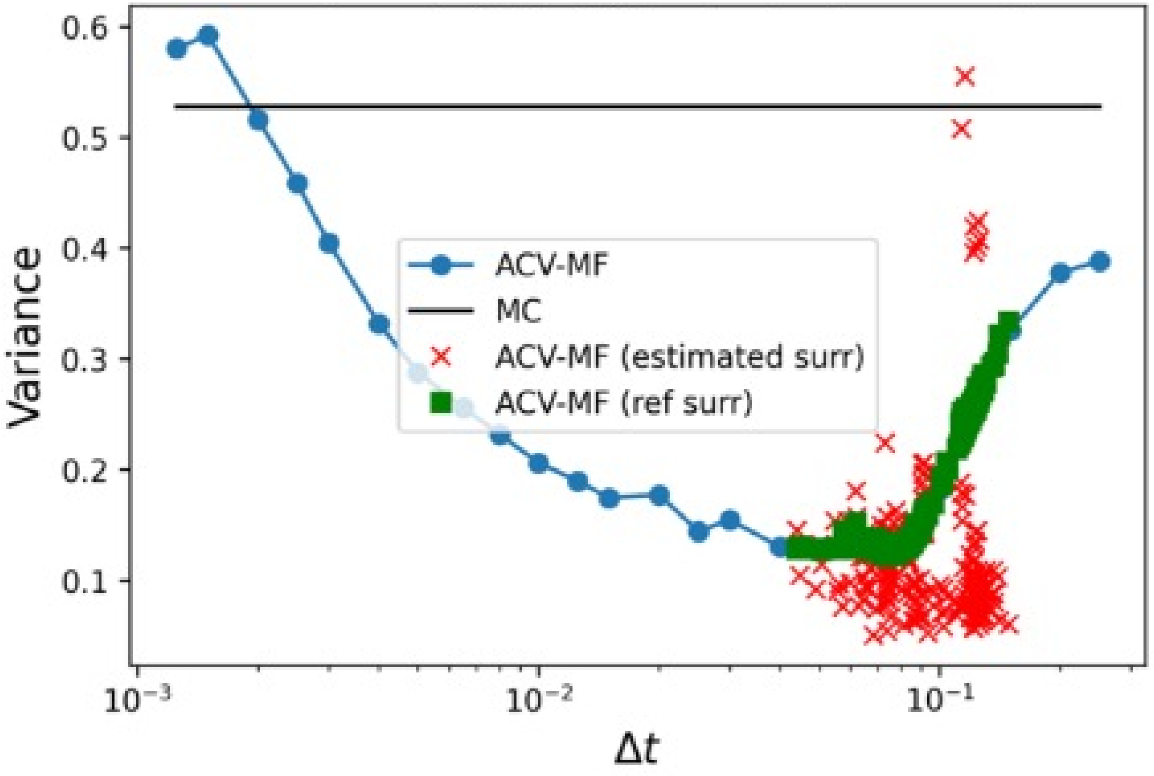


Smoothness is important for gradient-based optimization

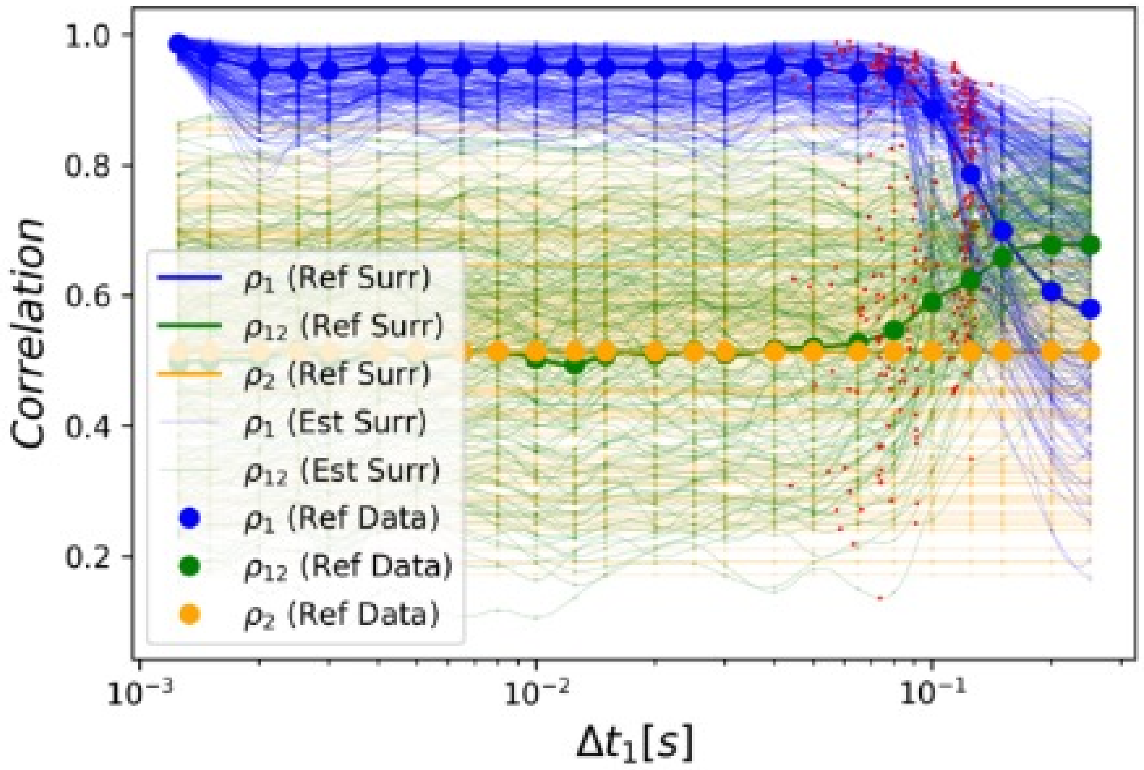
Results: Fewer Pilot Samples

- Optimization using $\mathrm{N_{tun}} = 24$, $\mathbf{N_{pilot} = 50}$ to build correlation surrogate models; 200 random trials

Estimator Variance vs. Time Step



Estimated Model Correlation vs. Time Step

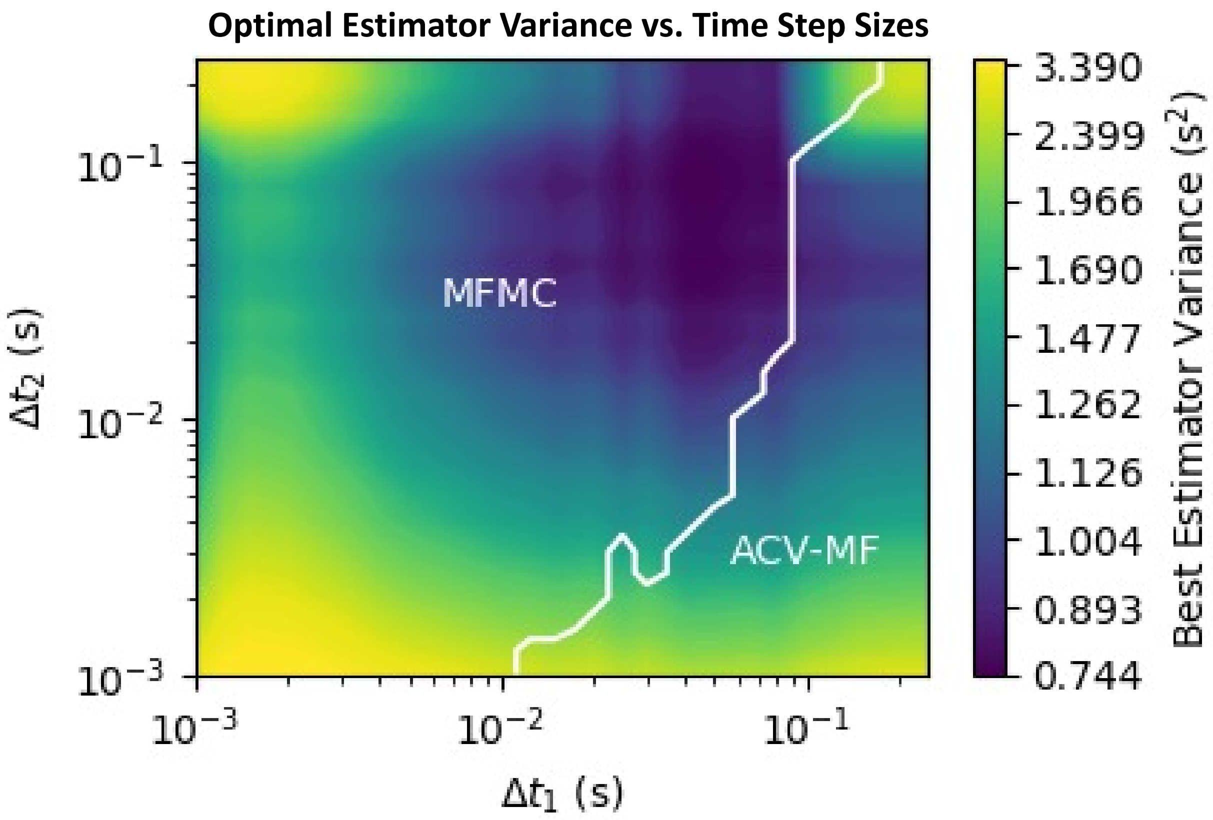


Statistical variability impacts optimization

Results: Two Tunable Lo-Fi Models

High-fidelity (Q): timestep = 0.001
Mid-fidelity (Q_1): timestep = $0.001 \leq \Delta t_1 \leq 0.25$
Low-fidelity (Q_2): timestep = 0.001 <= $\Delta t_2 \leq 0.25$

- Estimator variance was calculated directly for each $(\Delta t_1, \Delta t_2)$ pair using MFMC and ACV-MF and the optimal (minimum) value was plotted



Optimal ACV method (model graph) can be function of hyperparameters

Conclusions

- Model tuning can greatly affect estimator variance
- Optimization requires estimation (or knowledge) of correlations/costs as a function of tuning parameters
- Quality of the correlation surrogate is an important factor in tuning parameter optimization
- **Future Work:**
  - Automating sample allocation + model tuning optimization – implementation in Dakota
  - Global optimization with adaptive surrogate refinement
  - All-at-once optimization with model hierarchy

Questions?

- Email: james.e.warner@nasa.gov

- Reference:

    - G. F. Bomarito, G. Geraci, J. E. Warner, P. E. Leser, W. P. Leser, M. S. Eldred, J. D. Jakeman, A. A. Gorodetsky. *Improving Multi-Model Trajectory Simulation Estimators using Model Selection and Tuning*. 2022 AIAA Scitech Proceedings. January 2022. https://doi.org/10.2514/6.2022-1099.

- Acknowledgement:

    - *NASA researchers supported by the Entry Systems Modeling (ESM) project*
    - *Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.*