USNCCM 2023

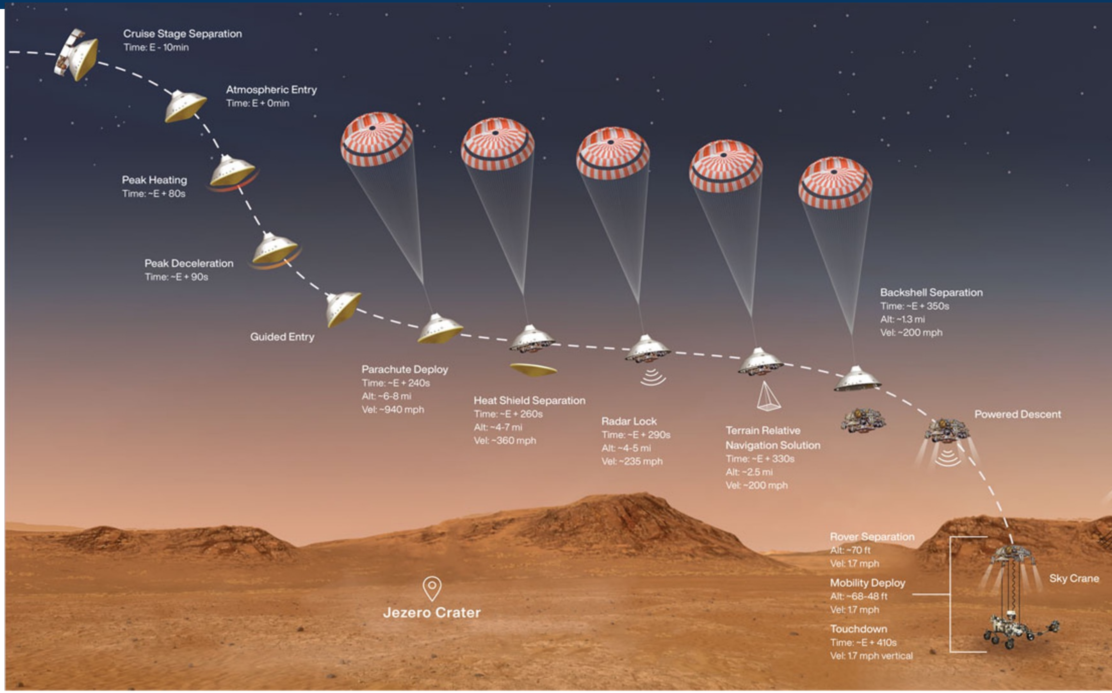
Multi-Output Approximate Control Variate Estimation for Enhanced Variance Reduction

Thomas Dixon & A. Gorodetsky, **University of Michigan**
G. Bomarito & J. Warner, **NASA Langley Research Center**

July 26, 2023

Uncertainty Quantification

- Example system
  - Entry, descent, and landing (EDL)
  - Perseverance Mars rover
- System uncertainties
  - Entry velocity
  - Entry location
- Multiple fidelity models
  - Full physics model
  - Reduced physics models
  - Machine learning
- Multiple outputs
  - Landing location
  - Landing velocity
- Multiple statistics
  - Mean
  - Variance
  - Sobol index

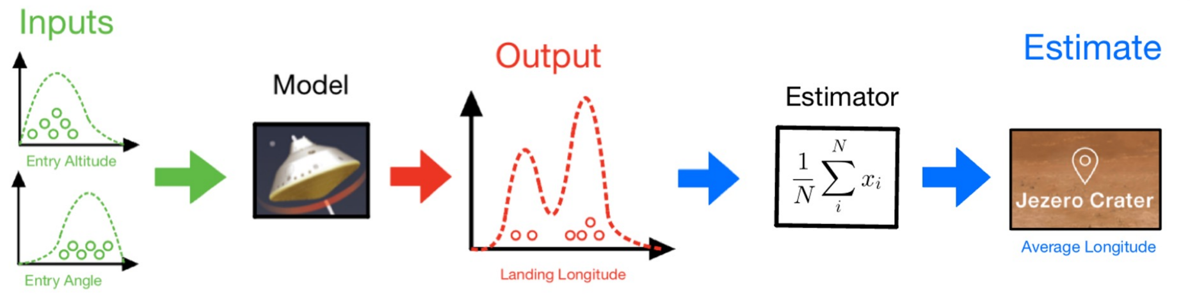


https://mars.nasa.gov/mars2020/timeline/landing/entry-descent-landing/

Thomas Dixon
tdixono@umich.edu

Sampling and Estimation

- Process:
  - Sample from the input space
  - Run the samples through the model
  - Estimate statistics using function evaluations
- Problem:
  - Accurate statistics are very expensive
  - Many function evaluations are required for accurate model statistics



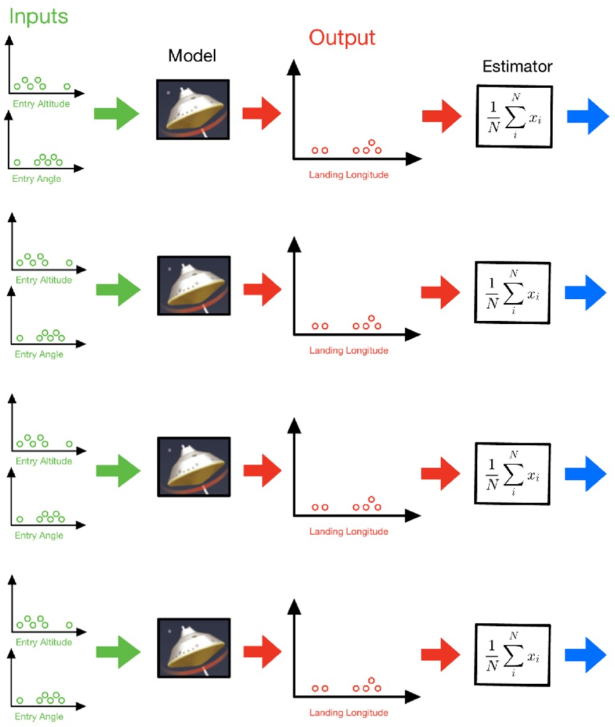
Thomas Dixon
tdixono@umich.edu

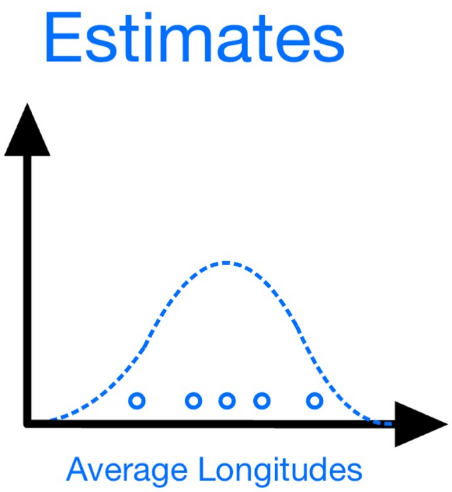
Variance Reduction

- Goal:
  - Reduce the variance of the estimator
  - Increase the confidence in an estimation
  - Reduce the cost of estimation for fixed confidence
- Monte Carlo Variance

Estimator Variance   Model Variance

$$\underline{\mathbb{V}ar[\hat{x}]} = \frac{1}{N}\underline{\mathbb{V}ar[x]}$$



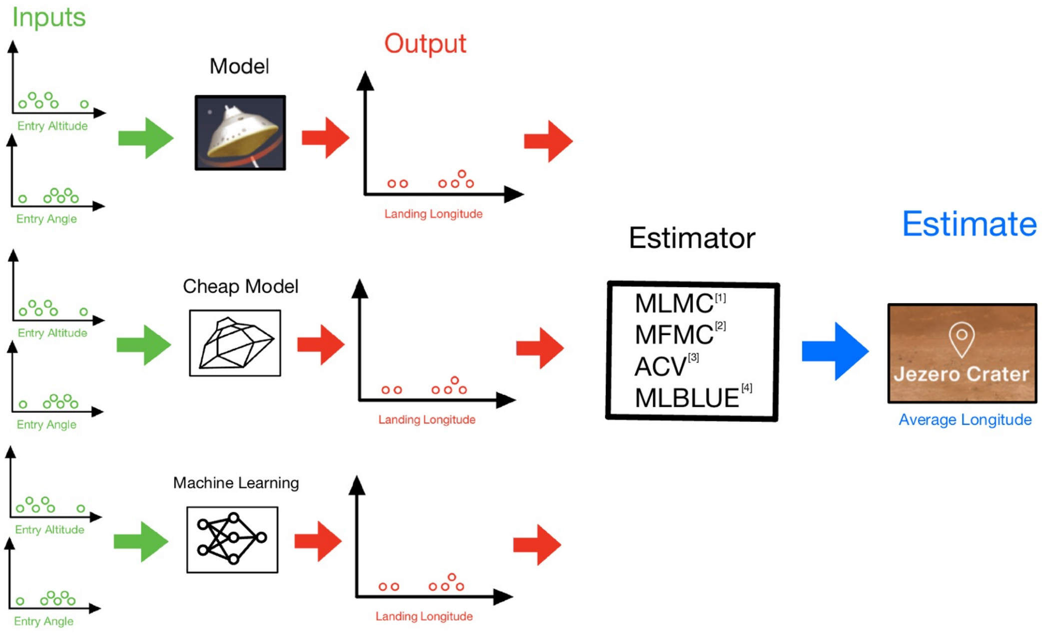


Thomas Dixon
tdixono@umich.edu

Multi-Fidelity Estimation

- Use multiple models for variance reduction
- Cheaper data is available from multiple sources

Every Method Requires: Covariance of Estimators



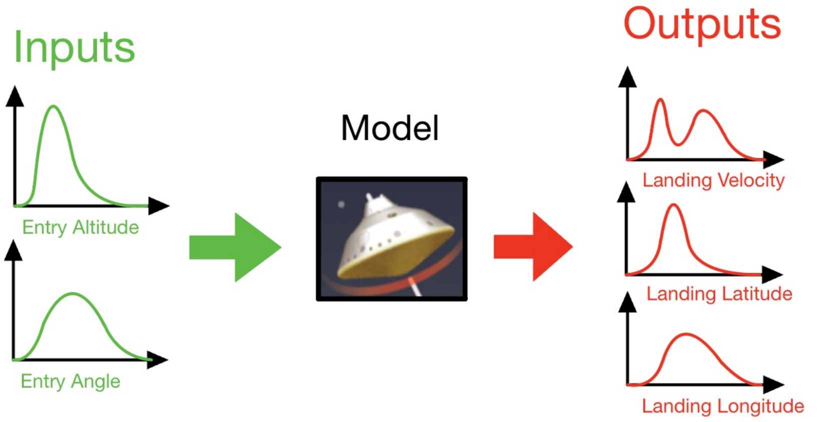
[1] Giles. (2007). “Variance Reduction through **Multilevel Monte Carlo** Path Calculations”
[2] Peherstorfer, et al. (2016). “Optimal Model Management for **Multifidelity Monte Carlo** Estimation”
[3] Gorodetsky, et al. (2020). “A generalized **approximate control variate** framework for multifidelity uncertainty quantification”
[4] Schaden, et al. (2020) “On **multilevel best linear unbiased estimators**”

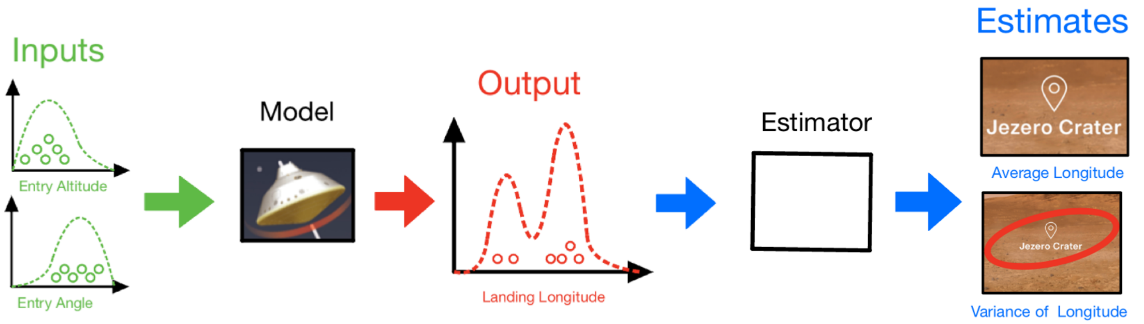
Thomas Dixon
tdixono@umich.edu

Multi-Output Estimation

- Multiple outputs of the model
  - Velocity, latitude, longitude
- Multiple statistics of the model
  - Simultaneous mean and variance estimation

- Previous Work:
- Multi-Output MLBLUE [6]





[5] Croci et al. (2023). "Multi-output multilevel best linear unbiased estimators via semidefinite programming"
[6] Destouches et al. (2023). "Multivariate extensions of the Multilevel Best Linear Unbiased Estimator for ensemble-variational data assimilation"

Thomas Dixon
tdixono@umich.edu

Contributions and Rest of Talk

- Contributions
  - Closed-form expressions for estimator covariances
  - Multi-output approximate control variate (ACV) estimator

- Outline
  - Control variates
    - Estimator covariances
  - Theoretical results
  - Experiments and results
    - EDL example
  - Future work

Control Variates

- Using correlated random variable to reduce variance

$$\tilde{\mathbf{Q}} = \mathbf{Q} + \alpha(\mathbf{C} - c)$$

High-fidelity **Low-fidelity**

Few samples **Many samples**

  - Control variate has known mean
  - Unbiased estimator
- Minimize variance to find optimal weight

$$\alpha^* = -\mathbb{C}ov[\mathbf{Q}, \mathbf{C}]\mathbb{V}ar[\mathbf{C}]^{-1}$$

Covariance of Estimators

- Optimal variance

  Variance is reduced

$$\mathbb{V}ar[\tilde{\mathbf{Q}}] = \mathbb{V}ar[\mathbf{Q}] - \mathbb{C}ov[\mathbf{Q}, \mathbf{C}]\mathbb{V}ar[\mathbf{C}]^{-1}\mathbb{C}ov[\mathbf{Q}, \mathbf{C}]^T$$

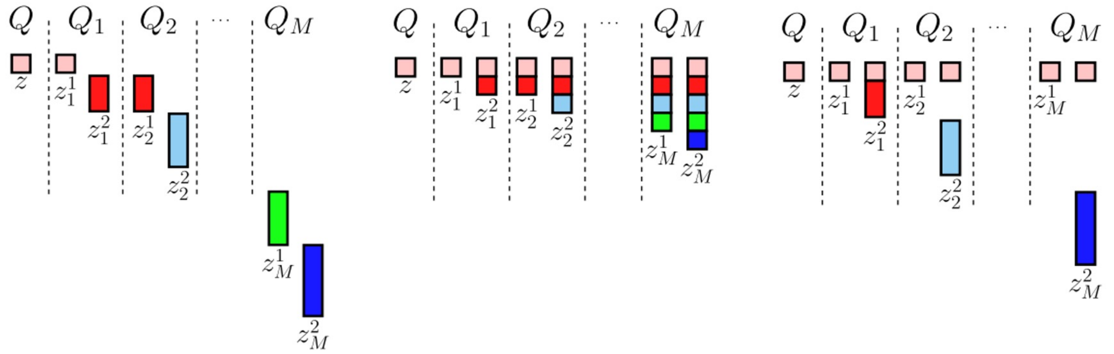
Thomas Dixon
tdixono@umich.edu

ACVs

- Control variate with unknown mean

$$\tilde{\mathbf{Q}} = \mathbf{Q}(\mathbf{z}) + \alpha(\mathbf{C}(\mathbf{z}_1) - \mathbf{C}(\mathbf{z}_2))$$

  - Need to **estimate the mean** of the control variate
  - Control variate estimators take different **input samples (z)**
- Multiple control variates



(a) Multilevel Monte Carlo (MLMC)

(b) Multifidelity Monte Carlo (MFMC)

(c) ACV Independent Samples

“A generalized approximate control variate framework for multifidelity uncertainty quantification.” Gorodetsky, et al. (2020)

Multi-Output ACV (MOACV)

- Creating one estimator for all outputs
  - Leverages correlations between model outputs
  - Uses model outputs across fidelities as control variates for other outputs
- Previously, separate estimators have been created for each model output
- **Old estimator equation**:
  - Two outputs
  - One low fidelity model

$$\boldsymbol{\Delta} = \begin{bmatrix} \boldsymbol{\Delta}_1 \\ \vdots \\ \boldsymbol{\Delta}_K \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1^1 - \mathbf{C}_1^2 \\ \vdots \\ \mathbf{C}_K^1 - \mathbf{C}_K^2 \end{bmatrix}$$

Unused correlation

ACV estimator

$$\left\{ \begin{bmatrix} \tilde{Q}_1 \\ \tilde{Q}_2 \end{bmatrix} = \begin{bmatrix} \hat{Q}_1 \\ \hat{Q}_2 \end{bmatrix} + \underbrace{\begin{bmatrix} \alpha_{1,1} & 0 \\ 0 & \alpha_{2,2} \end{bmatrix}}_{\text{Control variate weights}} \begin{bmatrix} \Delta^{(1)} \\ \Delta^{(2)} \end{bmatrix} \right.$$

First model output

Second model output

High-fidelity estimator

Thomas Dixon
tdixono@umich.edu

Multi-Output ACV (MOACV)

- Creating one estimator for all outputs
  - Leverages correlations between model outputs
  - Uses model outputs across fidelities as control variates for other outputs
- Previously, separate estimators have been created for each model output
- **Old estimator equation**:
  - Two outputs
  - One low fidelity model

$$\underline{\boldsymbol{\Delta}} = \begin{bmatrix} \boldsymbol{\Delta}_1 \\ \vdots \\ \boldsymbol{\Delta}_K \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1^1 - \mathbf{C}_1^2 \\ \vdots \\ \mathbf{C}_K^1 - \mathbf{C}_K^2 \end{bmatrix}$$

New weights

ACV estimator

$$\underbrace{\begin{bmatrix} \tilde{Q}_1 \\ \tilde{Q}_2 \end{bmatrix}}_{\text{ACV estimator}} = \underbrace{\begin{bmatrix} \hat{Q}_1 \\ \hat{Q}_2 \end{bmatrix}}_{\text{High-fidelity estimator}} + \underbrace{\begin{bmatrix} \alpha_{1,1} & \boxed{\alpha_{1,2}} \\ \boxed{\alpha_{2,1}} & \alpha_{2,2} \end{bmatrix}}_{\text{Control variate weights}} \begin{bmatrix} \Delta^{(1)} \\ \Delta^{(2)} \end{bmatrix}$$

First model output

Second model output

High-fidelity estimator

Control variate weights

Thomas Dixon
tdixono@umich.edu

Combining Estimators - Multiple Statistics

- Simultaneous statistic estimation
  - Further reduced variance
- Correlations between estimators can also be extracted
- Equation
  - Any number of outputs
    - Mean
    - Variance
  - One low-fidelity model
  - Boxed in green are the new control variate weights

$$\begin{array}{r} \text{ACV mean estimator} \rightarrow \\ \text{ACV variance estimator} \rightarrow \end{array} \begin{bmatrix} \tilde{Q}_\mu \\ \tilde{Q}_V \end{bmatrix} = \begin{bmatrix} \hat{Q}_\mu \\ \hat{Q}_V \end{bmatrix} + \begin{bmatrix} \alpha_{\mu,\mu} & \boxed{\alpha_{\mu,V}} \\ \boxed{\alpha_{V,\mu}} & \alpha_{V,V} \end{bmatrix} \begin{bmatrix} \Delta_{1,\mu} \\ \Delta_{1,V} \end{bmatrix}$$

Thomas Dixon
tdixono@umich.edu

Types of Estimators

Estimator	Statistics	Quantities of Interest	Fidelities
MC - Monte Carlo	Single	Single	Single
ACV	Single	Single	**Multiple**
MOACV	Single	**Multiple**	**Multiple**
Combined MOACV	**Multiple**	**Multiple**	**Multiple**

Theoretical Results
Mean Estimator Covariances

- Mean estimation
- Covariance between mean estimators
- Required covariances

$$\alpha^* = -\mathbb{C}ov[\mathbf{Q}, \mathbf{\Delta}]\mathbb{V}ar[\mathbf{\Delta}]^{-1}$$

$$\hat{\mathbf{Q}}_i^{\mathcal{N}} = \frac{1}{N}\sum_{t}^{N} f_i(n^{(t)})$$

$$\mathbb{C}ov[\hat{\mathbf{Q}}_i^{\mathcal{N}}, \hat{\mathbf{Q}}_j^{\mathcal{M}}] = \frac{|\mathcal{N} \cap \mathcal{M}|}{NM}\mathbb{C}ov\left[f_i, f_j\right]$$

$$\mathbb{C}ov[\mathbf{Q}, \mathbf{\Delta}] = \mathbf{G} \circ \mathbb{C}ov[f_0, \mathbf{f}]$$

$$G_i = \frac{L_{0,i^1}}{L_0 L_{i^1}} - \frac{L_{0,i^2}}{L_0 L_{i^2}}$$

$$\mathbf{G} = G \otimes \mathbf{1}_{D\times D}$$

$$\mathbb{V}ar[\mathbf{\Delta}] = \mathbf{F} \circ \mathbb{V}ar[\mathbf{f}]$$

$$F_{i,j} = \begin{cases} \frac{1}{L_{i^1}} + \frac{1}{L_{i^2}} - 2\frac{L_{i^1,i^2}}{L_{i^1}L_{i^2}} & \textit{if } i{=}j \\ \frac{L_{i^1,j^1}}{L_{i^1}L_{j^1}} - \frac{L_{i^1,j^2}}{L_{i^1}L_{j^2}} - \frac{L_{i^2,j^1}}{L_{i^2}L_{j^1}} + \frac{L_{i^2,j^2}}{L_{i^2}L_{j^2}} & \textit{otherwise} \end{cases}$$

$$\mathbf{F} = F \otimes \mathbf{1}_{D\times D}$$

Theoretical Results

Variance Estimator Covariances

- Variance estimation
- Covariance between variance estimators

$$\hat{\mathbf{Q}}_i^{\mathcal{N}} = \frac{1}{2N(N-1)} \sum_s^N \sum_t^N \left(f_i(n_s) - f_i(n_t)\right)\left(f_i(n_s) - f_i(n_t)\right)^T$$

$$\mathbb{C}ov[\hat{\mathbf{Q}}_i^{\mathcal{N}}, \hat{\mathbf{Q}}_j^{\mathcal{M}}] = \frac{P(P-1)}{N(N-1)M(M-1)} \left[\mathbb{C}ov[f_i, f_j]^{\otimes 2} + \left(\mathbf{1}^T \otimes \mathbb{C}ov[f_i, f_j] \otimes \mathbf{1}\right) \circ \left(\mathbf{1} \otimes \mathbb{C}ov[f_i, f_j] \otimes \mathbf{1}^T\right)\right]$$

$$+ \frac{P}{NM} \mathbb{C}ov\left[(f_i - \mathbb{E}[f_i])^{\otimes 2}, (f_j - \mathbb{E}[f_j])^{\otimes 2}\right]$$

Thomas Dixon
tdixono@umich.edu

Theoretical Results

Combined Estimator Covariances

- Mean and variance estimation
- Covariance between estimators

  Using previous results

$$\hat{\mathbf{Q}}_i^{\mathcal{N}} = \begin{bmatrix} \hat{\boldsymbol{\mu}}_i^{\mathcal{N}} \\ \hat{\mathbf{V}}_i^{\mathcal{N}} \end{bmatrix} = \begin{bmatrix} \frac{1}{N}\sum_t^N f_i(n_t) \\ \frac{1}{2N(N-1)} \sum_s^N \sum_t^N (f_i(n_s) - f_i(n_t))^{\otimes 2} \end{bmatrix}$$

$$\mathbb{C}ov[\hat{\mathbf{Q}}_i^{\mathcal{N}}, \hat{\mathbf{Q}}_j^{\mathcal{M}}] = \begin{bmatrix} \mathbb{C}ov[\hat{\boldsymbol{\mu}}_i^{\mathcal{N}}, \hat{\boldsymbol{\mu}}_j^{\mathcal{M}}] & \mathbb{C}ov[\hat{\boldsymbol{\mu}}_i^{\mathcal{N}}, \hat{\mathbf{V}}_j^{\mathcal{M}}] \\ \mathbb{C}ov[\hat{\mathbf{V}}_i^{\mathcal{N}}, \hat{\boldsymbol{\mu}}_j^{\mathcal{M}}] & \mathbb{C}ov[\hat{\mathbf{V}}_i^{\mathcal{N}}, \hat{\mathbf{V}}_j^{\mathcal{M}}] \end{bmatrix}$$

$$\mathbb{C}ov[\hat{\boldsymbol{\mu}}_i^{\mathcal{N}}, \hat{\mathbf{V}}_j^{\mathcal{M}}] = \frac{P}{NM}\mathbb{C}ov[f_i, (f_j - \mathbb{E}[f_j])^{\otimes 2}]$$

Thomas Dixon
tdixono@umich.edu

Theoretical Results

Combined Estimator Covariances

- Mean and variance estimation
- Covariance between estimators

  Using previous results

$$\hat{\mathbf{Q}}_i^{\mathcal{N}} = \begin{bmatrix} \hat{\mu}_i^{\mathcal{N}} \\ \hat{\mathbf{V}}_i^{\mathcal{N}} \end{bmatrix} = \begin{bmatrix} \frac{1}{N}\sum_t^N f_i(n_t) \\ \frac{1}{2N(N-1)}\sum_s^N \sum_t^N (f_i(n_s) - f_i(n_t))^{\otimes 2} \end{bmatrix}$$

$$\mathbb{C}ov[\hat{\mathbf{Q}}_i^{\mathcal{N}}, \hat{\mathbf{Q}}_j^{\mathcal{M}}] = \begin{bmatrix} \mathbb{C}ov[\hat{\mu}_i^{\mathcal{N}}, \hat{\mu}_j^{\mathcal{M}}] & \mathbb{C}ov[\hat{\mu}_i^{\mathcal{N}}, \hat{\mathbf{V}}_j^{\mathcal{M}}] \\ \mathbb{C}ov[\hat{\mathbf{V}}_i^{\mathcal{N}}, \hat{\mu}_j^{\mathcal{M}}] & \mathbb{C}ov[\hat{\mathbf{V}}_i^{\mathcal{N}}, \hat{\mathbf{V}}_j^{\mathcal{M}}] \end{bmatrix}$$

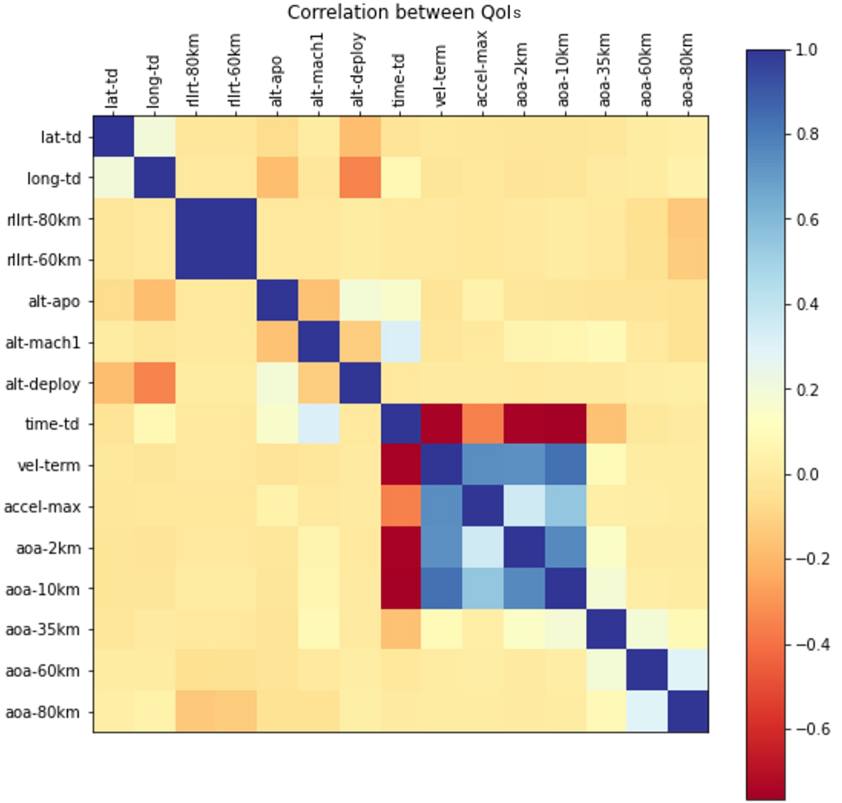
$$\mathbb{C}ov[\hat{\mu}_i^{\mathcal{N}}, \hat{\mathbf{V}}_j^{\mathcal{M}}] = \frac{P}{NM}\mathbb{C}ov[f_i, (f_j - \mathbb{E}[f_j])^{\otimes 2}]$$

Main takeaway:
Closed-form expressions for the covariance between estimators

Thomas Dixon
tdixono@umich.edu

EDL Application

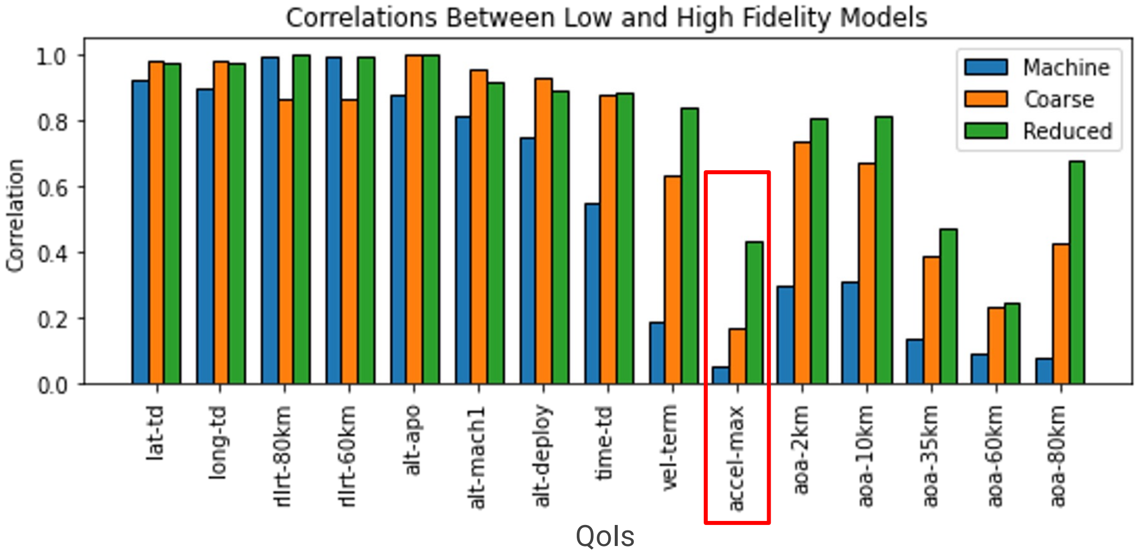
- "Multi-Model Monte Carlo Estimators for Trajectory Simulation" by Warner, et al. (2021)
- 75 uncertain inputs
- 15 Quantities of Interest (QoIs)
- One high-fidelity model
  - Trajectory simulation
- Three low-fidelity models
  - Reduced physics model
  - Coarse time step model
  - Machine learning model
- Mean and variance estimation
  - Optimized sample allocation
  - Minimized weighted sum of variances



Thomas Dixon
tdixono@umich.edu

Correlation Between Models

- Plotting correlations between high-fidelity and the low-fidelity models
  - Each QoI has a separate correlation between models
- Maximum acceleration has low correlations between models



Thomas Dixon
tdixono@umich.edu

Types of Estimators

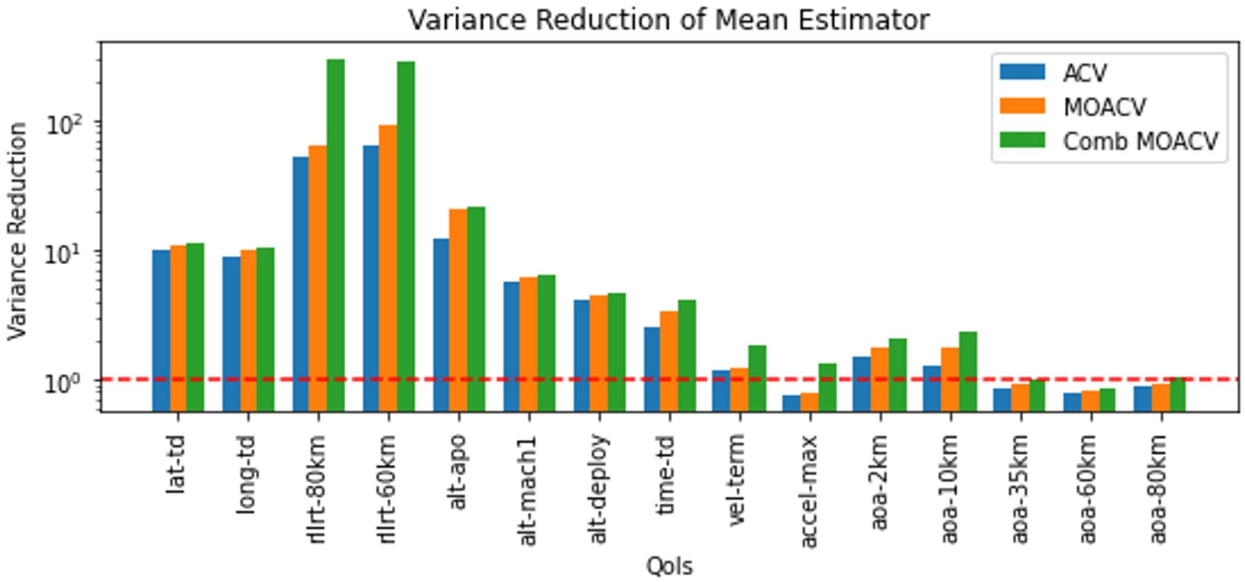
Estimator	Statistics	QoIs	Fidelities
MC	Single	Single	Single
ACV	Single	Single	**Multiple**
MOACV	Single	**Multiple**	**Multiple**
Combined MOACV	**Multiple**	**Multiple**	**Multiple**

Thomas Dixon
tdixono@umich.edu

Results - Mean Estimation

- Variance reduction for each QoI
- Significant reduction in QoI with high correlations
- MC achieved at red line

$$\text{Reduction} = \frac{\mathbb{V}ar[\tilde{Z}]_{MC}}{\mathbb{V}ar[\tilde{Z}]_{CV}}$$

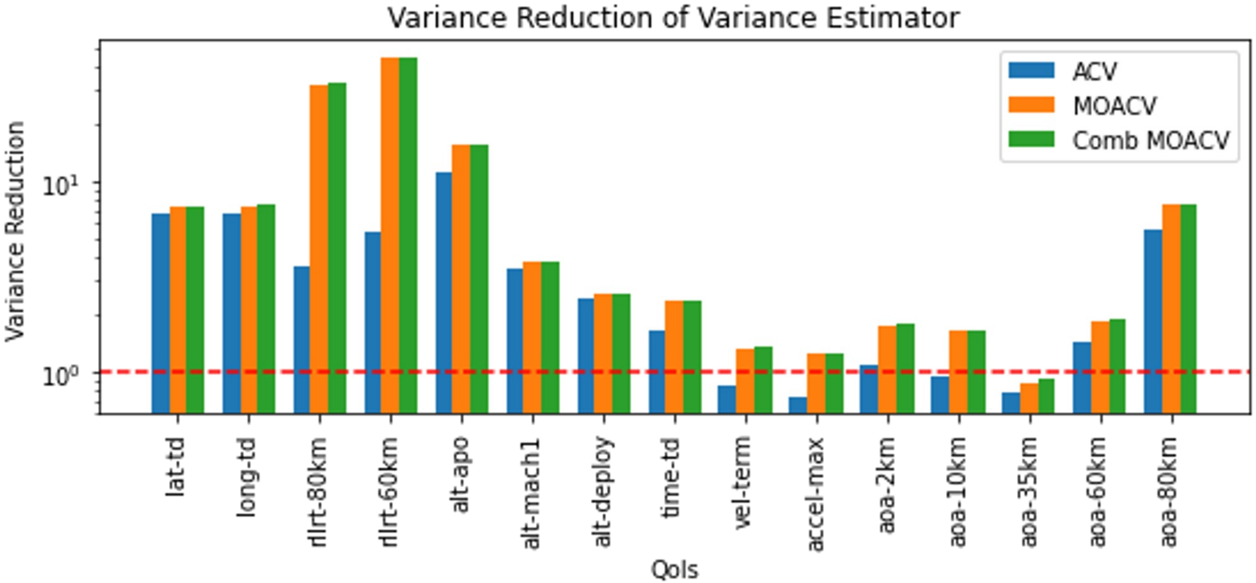


Thomas Dixon
tdixono@umich.edu

Results - Variance Estimation

- Significant variance reduction achieved compared to traditional approaches

$$\text{Reduction} = \frac{\mathbb{V}ar[\tilde{Z}]_{MC}}{\mathbb{V}ar[\tilde{Z}]_{CV}}$$



Thomas Dixon
tdixono@umich.edu

Future Work

Stochastic Optimization

- Goal: Minimize finite-sum objective function
- Stochastic gradient descent
- Approximate the gradient of the objective

$$\min_{d \in \mathbf{D}} \quad J(d,z) = \min_{d \in \mathbf{D}} \quad \frac{1}{N} \sum_{z=1}^{N} \mathbf{q}(d,z)$$

$$\mathbf{g}(d) = \nabla_d J(d)$$

$$d_{k+1} = d_k + \gamma \mathbf{g}(d) \quad \rightarrow \quad d_{k+1} = d_k + \gamma \hat{\mathbf{g}}(d)$$

- Variance reduction methods
    - Variance of estimate approaches 0
    - Examples:
        - Stochastic variance reduced gradient (SVRG)[1]
        - Stochastic average gradient (SAG)[2]

$$\mathbb{V}ar[\hat{\mathbf{g}}(d)] \rightarrow 0$$

- Use MOACV to reduce variance further

[1] Johnson, et al. (2013). “Accelerating stochastic gradient descent using predictive variance reduction”
[2] Roux, et al. (2012). “A stochastic gradient method with an exponential convergence rate for finite training sets”

Thomas Dixon
tdixono@umich.edu

Conclusion

- Derived closed-form expressions for estimator covariance
- New multi-output estimator outperforms previous methods

- **Questions?**

Thomas Dixon
tdixono@umich.edu

Supplementary Material