-9108C

National Laboratories

Multi-Output Surrogate Construction for Fusion Simulations

PRESENTED BY

Kathryn Maupin, Anh Tran, Michael Glinsky

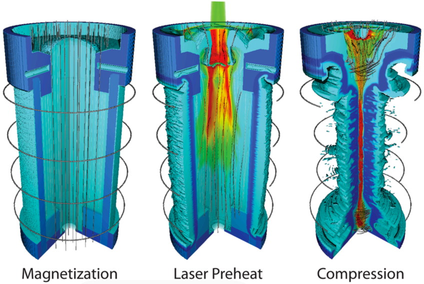
USNCCM

July 29, 2021

Motivation: MagLIF

Magnetized Linear Inertial Fusion relies on compression of a magnetized, laser-heated fuel to achieve thermonuclear ignition

- Experiments performed on Sandia's Z Pulsed Power Facility
- Preheated deuterium fuel
- Solid beryllium liner



P.F. Knapp, *et al*, 2015

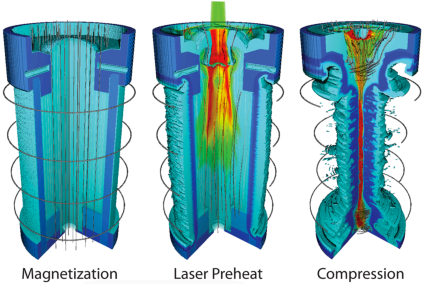
Motivation: MagLIF

The state of the fuel is not directly observable

Physicists rely on diagnostic metrics to understand:

- Target performance
- Impact of modifications
- Importance of sources of degradation

The calibration of these diagnostics becomes a multi-objective inference problem



P.F. Knapp, *et al*, 2015

Calibration

Bayesian calibration naturally incorporates uncertainties during calibration and prediction

$$\pi(\boldsymbol{\theta}|\mathbf{d}) = \frac{\pi(\mathbf{d}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{d})}$$

- Both the prediction and propagation phases require many runs of the model and incur significant computational expense

Calibration

Bayesian calibration naturally incorporates uncertainties during calibration and prediction

$$\pi(\boldsymbol{\theta}|\mathbf{d}) = \frac{\pi(\mathbf{d}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{d})}$$

- Both the prediction and propagation phases require many runs of the model and incur significant computational expense $\rightarrow$ Surrogate

Calibration

Bayesian calibration naturally incorporates uncertainties during calibration and prediction

$$\pi(\boldsymbol{\theta}|\mathbf{d}) = \frac{\pi(\mathbf{d}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{d})}$$

- Both the prediction and propagation phases require many runs of the model and incur significant computational expense $\rightarrow$ Surrogate

The current standard of practice for multi-output surrogate models is to use individual phenomenological surrogates

- Correlation between QoIs is lost

Calibration

Bayesian calibration naturally incorporates uncertainties during calibration and prediction

$$\pi(\boldsymbol{\theta}|\mathbf{d}) = \frac{\pi(\mathbf{d}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{d})}$$

- Both the prediction and propagation phases require many runs of the model and incur significant computational expense $\rightarrow$ Surrogate

The current standard of practice for multi-output surrogate models is to use individual phenomenological surrogates

- Correlation between QoIs is lost $\rightarrow$ Co-prediction

Calibration

Bayesian calibration naturally incorporates uncertainties during calibration and prediction

$$\pi(\boldsymbol{\theta}|\mathbf{d}) = \frac{\pi(\mathbf{d}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{d})}$$

- Both the prediction and propagation phases require many runs of the model and incur significant computational expense $\rightarrow$ Surrogate

The current standard of practice for multi-output surrogate models is to use individual phenomenological surrogates

- Correlation between QoIs is lost $\rightarrow$ Co-prediction

Goal: Construct co-predictive surrogate model

Gaussian Processes

A Gaussian process is a stochastic process such that every finite collection of its random variables has a multivariate normal distribution

$$f(\mathbf{x}) \sim \mathcal{N}(\mu(\mathbf{x}), \Sigma(\mathbf{x}, \mathbf{x}'))$$

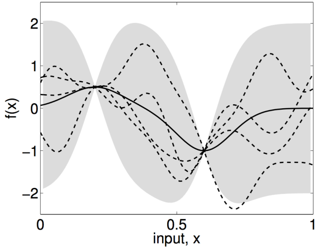
Σ is completely defined by the correlation function $k(\mathbf{x}, \mathbf{x}')$

- Squared exponential
- Matern

GP surrogates interpolate data points and provide uncertainty estimates for each output value

Computation of prediction mean and variance requires inversion of the $N \times N$ correlation matrix

$$\mathbf{R}_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j)$$



C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning*

Multi-Output Gaussian Processes (MOGP)

Consider the multi-output vector

$$\mathbf{f} = [f_1, \ldots, f_T]^T$$

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{N}(\boldsymbol{\mu}(\mathbf{x}), \Sigma(\mathbf{x}, \mathbf{x}'))$$

Σ is defined by a multi-output covariance $K(\mathbf{x}, \mathbf{x}')$

$$K(\mathbf{x}, \mathbf{x}') = \begin{bmatrix} k_{11}(\mathbf{x}, \mathbf{x}') & \ldots & k_{1T}(\mathbf{x}, \mathbf{x}') \\ \vdots & \ddots & \vdots \\ k_{T1}(\mathbf{x}, \mathbf{x}') & \ldots & k_{TT}(\mathbf{x}, \mathbf{x}') \end{bmatrix}$$

Computation of prediction mean and variance requires inversion of the $NT \times NT$ correlation matrix

$$\begin{aligned} \mathbf{R}_{IJ} &= \mathbf{K}_{IJ} \\ [\mathbf{K}_{IJ}]_{ij} &= k_{IJ}(\mathbf{x}_i, \mathbf{x}_j) \end{aligned}$$

Linear Model of Coregionalization (LMC)

Define Q covariance functions $k_q(\mathbf{x}, \mathbf{x}')$ and sample R_q latent functions

$$u_q^i \sim \mathcal{GP}(0, k_q(\mathbf{x}, \mathbf{x}'))$$

For output t,

$$f_t(\mathbf{x}) = \sum_{q=1}^{Q} \sum_{i=1}^{R_q} a_{t,q}^i u_q^i(\mathbf{x})$$

The cross-covariance is given by

$$cov[\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')] = \sum_{q=1}^{Q} \mathbf{A}_q \mathbf{A}_q^T k_q(\mathbf{x}, \mathbf{x}') = \sum_{q=1}^{Q} \mathbf{B}_q k_q(\mathbf{x}, \mathbf{x}')$$

where $\mathbf{A}_q = [\mathbf{a}_q^1\, \mathbf{a}_q^2\, \dots\, \mathbf{a}_q^{R_q}]$

Two special cases

- $Q = 1 \Rightarrow$ intrinsic coregionalization model (ICM)
- $R_q = 1 \Rightarrow$ semi-parametric latent factor model (SLFM)

	LMC	ICM	SLFM
$f_t(x) =$	$\sum_{q=1}^{Q} \sum_{i=1}^{R_q} a_{t,q}^i u_q^i(\mathbf{x})$	$\sum_{i=1}^{R} a_t^i u^i(\mathbf{x})$	$\sum_{q=1}^{Q} a_{t,q} u_q(\mathbf{x})$
$cov[\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')] =$	$\sum_{q=1}^{Q} \mathbf{B}_q k_q(\mathbf{x}, \mathbf{x}')$	$\mathbf{B} k(\mathbf{x}, \mathbf{x}')$	$\sum_{q=1}^{Q} \mathbf{B}_q k_q(\mathbf{x}, \mathbf{x}')$
$\mathbf{A}_{(q)} =$	$[\mathbf{a}_q^1\, \mathbf{a}_q^2\, \ldots\, \mathbf{a}_q^{R_q}]$	$[\mathbf{a}^1\, \mathbf{a}^2\, \ldots\, \mathbf{a}^R]$	$\mathbf{a}_q$

Considerations

- k_q can be the same function with different hyperparameters, or different function types
- Larger Q increases flexibility (up to $Q = T$), but with computational cost

Benchmarking examples

Forrester

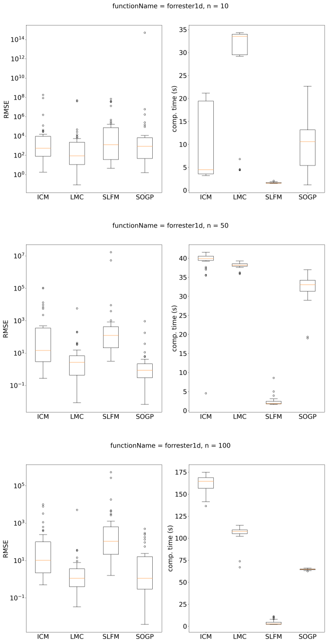
- 1 parameter
- $\rho = 0.71$

Accuracy:

- SLFM performs the worst, particularly as the number of parameters increases
- ICM and LMC are competitive with SOGP

Expense:

- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points



Benchmarking examples

Branin

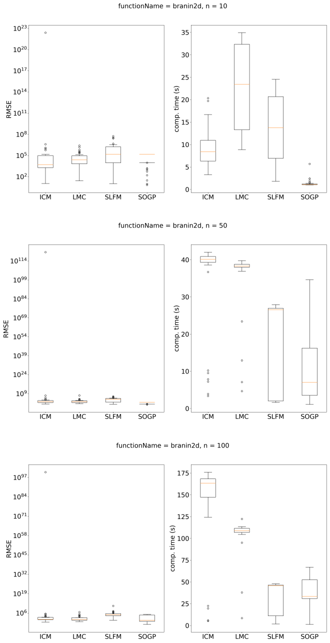
- 2 parameters
- $\rho = 0.68$

Accuracy:

- SLFM performs the worst, particularly as the number of parameters increases
- ICM and LMC are competitive with SOGP

Expense:

- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points



Benchmarking examples

Dette & Pepelyshev

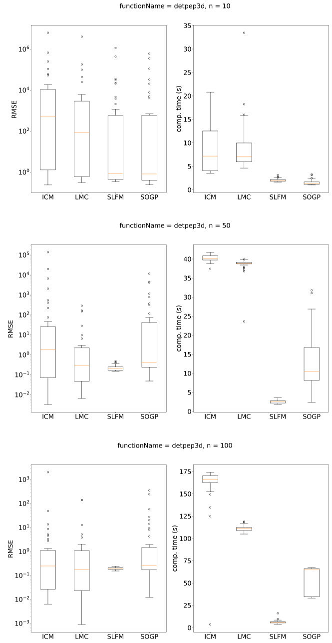
- 3 parameters
- $\rho = 0.68$

Accuracy:

- SLFM performs the worst, particularly as the number of parameters increases
- ICM and LMC are competitive with SOGP

Expense:

- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points



Benchmarking examples

Gramacy & Lee

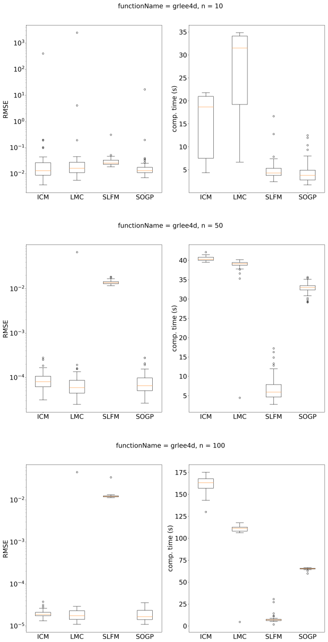
- 4 parameters
- $\rho = 0.83$

Accuracy:

- SLFM performs the worst, particularly as the number of parameters increases
- ICM and LMC are competitive with SOGP

Expense:

- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points



Benchmarking examples

Friedman

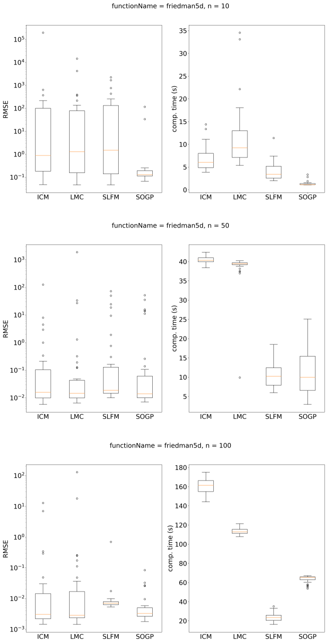
- 5 parameters
- $\rho = 0.98$

Accuracy:

- SLFM performs the worst, particularly as the number of parameters increases
- ICM and LMC are competitive with SOGP

Expense:

- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points



Benchmarking examples

Borehole

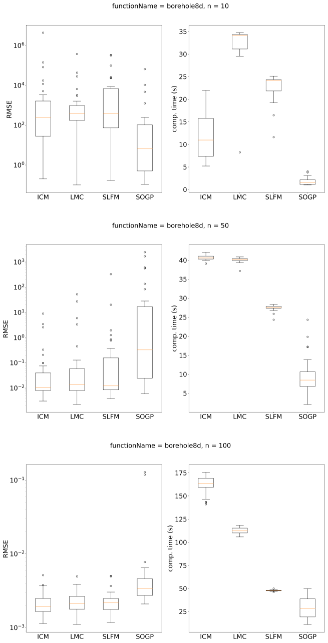
- 8 parameters
- $\rho = 1$

Accuracy:

- SLFM performs the worst, particularly as the number of parameters increases
- ICM and LMC are competitive with SOGP

Expense:

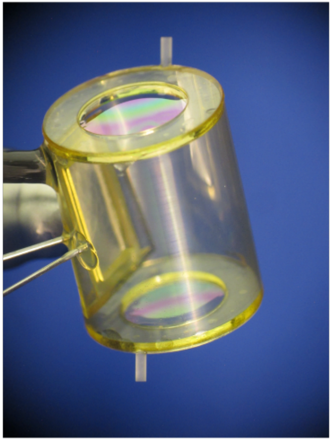
- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points

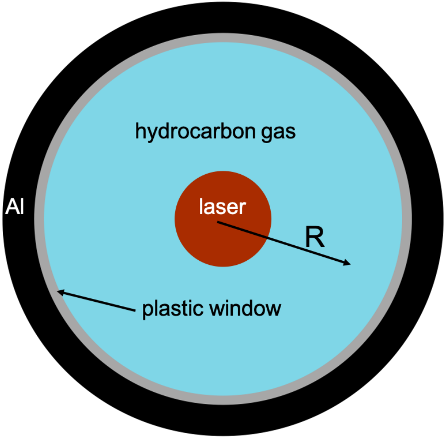


VISAR Experiment Simulations

VISAR = Velocity Interferometer System for Any Reflector

- Measures the shocks that occur during NIF experiments

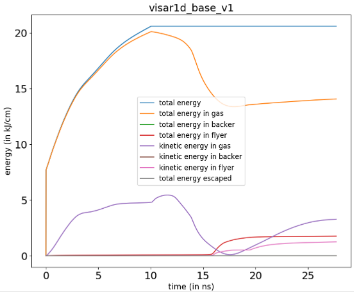


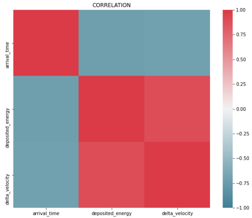


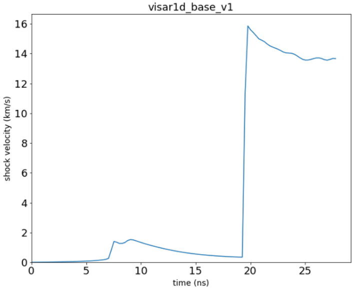
VISAR Experiment Simulations

1D simulation using Hydra

- Input parameters:
  - Deposition radius $\sim \mathcal{U}[400\mu m, 1200\mu m]$
  - Deposition temperature $\sim \mathcal{U}[0.8keV, 2.2keV]$
  - Deposition time $\sim \mathcal{U}[5ns, 15ns]$
- Outputs
  - Deposited energy
  - Arrival time of main shock
  - Delta velocity of main shock







VISAR Experiment Simulations

Computation time comparisons are similar to benchmark examples

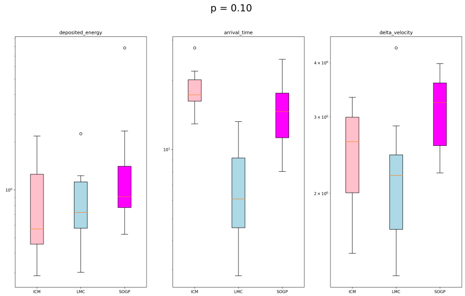
- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points

LMC performs the best

- ICM is hit or miss, but better with fewer build points

340 data points

- $p =$ percentage of points used



VISAR Experiment Simulations

Computation time comparisons are similar to benchmark examples

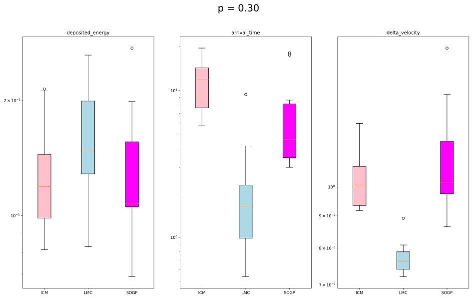
- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points

LMC performs the best

- ICM is hit or miss, but better with fewer build points

340 data points

- $p =$ percentage of points used



VISAR Experiment Simulations

Computation time comparisons are similar to benchmark examples

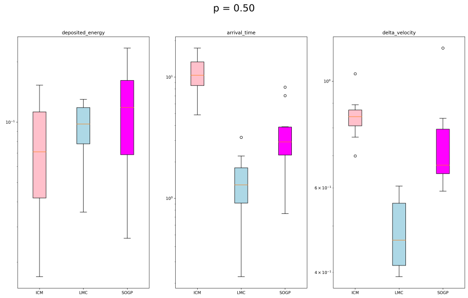
- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points

LMC performs the best

- ICM is hit or miss, but better with fewer build points

340 data points

- p = percentage of points used



VISAR Experiment Simulations

Computation time comparisons are similar to benchmark examples

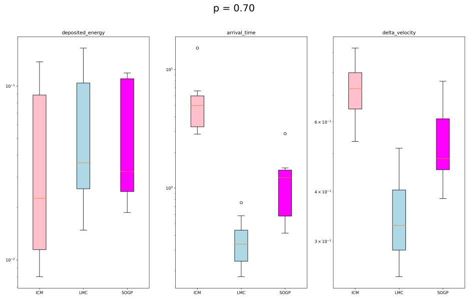
- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points

LMC performs the best

- ICM is hit or miss, but better with fewer build points

340 data points

- $p =$ percentage of points used



VISAR Experiment Simulations

Computation time comparisons are similar to benchmark examples

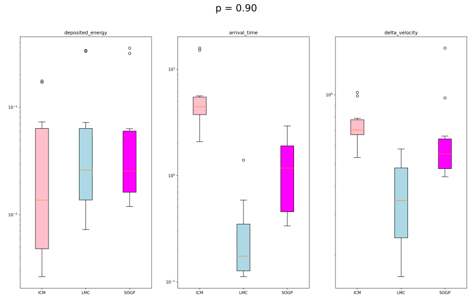
- SOGP is cheaper than ICM and LMC
- ICM is less expensive than LMC with fewer build points

LMC performs the best

- ICM is hit or miss, but better with fewer build points

340 data points

- $p =$ percentage of points used



Conclusions

Presented examples to compare three methods for calibrating MOGPs

- ICM and LMC are favorable over SLFM
- Benchmark examples: mixed results
- VISAR example: LMC outperforms SOGP

Next steps:

- Extend methodology to “field” data
- Include physics constraints
- Incorporate information from causal statistics

Thank you

Questions?