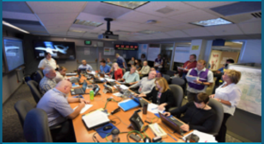
11009C

National Laboratories

Informing Missing Physics with Model Form Error and Model Selection



PRESENTED BY

Kathryn Maupin, Jaideep Ray, Anh Tran, and Teresa Portone

USACM UQ-MLIP

18-19 August 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

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

Calibration of Computer Models

Experimental data $=$ Model output $+$ error

$$d(x) = m(\theta, x) + \varepsilon$$

Significant disagreement between the data and the model can be due to model form error (model discrepancy, model inadequacy, structural error)

Calibration of Computer Models

Experimental data $=$ Model output $+$ error

$$d(x) = m(\theta, x) + \delta(x) + \varepsilon$$

Significant disagreement between the data and the model can be due to model form error (model discrepancy, model inadequacy, structural error)

Calibration of Computer Models

Experimental data = Model output + error

$$d(x) = m(\theta, x) + \delta(x) + \varepsilon$$

Significant disagreement between the data and the model can be due to model form error (model discrepancy, model inadequacy, structural error)

$\Rightarrow$ Something is missing from the model

- Incomplete understanding of physics
- Use of low-fidelity or reduced-order model

Calibration of Computer Models

Experimental data $=$ Model output $+$ error

$$d(x) = m(\theta, x) + \delta(x) + \varepsilon$$

Significant disagreement between the data and the model can be due to model form error (model discrepancy, model inadequacy, structural error)

$\Rightarrow$ Something is missing from the model

- Incomplete understanding of physics
- Use of low-fidelity or reduced-order model

$\Rightarrow$ **What can $\delta(x)$ tell us about how to improve $m(\theta)$?**

Calibration of Model Parameters

Calibrate model parameters θ via Bayes' Rule

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

- $\pi(\theta)$ captures any prior information that is known about the parameters before calibration
- $\pi(d|\theta)$ is the likelihood that the parameters will be able to reproduce the data
- $\pi(d)$ is the model evidence
- $\pi(\theta|d)$ is the updated distribution of parameter values

Calibration of Model Parameters

Calibrate model parameters θ via Bayes' Rule

$$\pi(\theta|d) = \frac{\pi(d|\theta)\overbrace{\pi(\theta)}^{\text{Prior}}}{\pi(d)}$$

- $\pi(\theta)$ captures any prior information that is known about the parameters before calibration
- $\pi(d|\theta)$ is the likelihood that the parameters will be able to reproduce the data
- $\pi(d)$ is the model evidence
- $\pi(\theta|d)$ is the updated distribution of parameter values

Calibration of Model Parameters

Calibrate model parameters θ via Bayes' Rule

$$\pi(\theta|d) = \frac{\overbrace{\pi(d|\theta)}^{\text{Likelihood}}\pi(\theta)}{\pi(d)}$$

- $\pi(\theta)$ captures any prior information that is known about the parameters before calibration
- $\pi(d|\theta)$ is the likelihood that the parameters will be able to reproduce the data
- $\pi(d)$ is the model evidence
- $\pi(\theta|d)$ is the updated distribution of parameter values

Calibration of Model Parameters

Calibrate model parameters θ via Bayes' Rule

$$\pi(\theta|d) = \frac{\pi(d|\theta)\pi(\theta)}{\underbrace{\pi(d)}_{\text{Evidence}}}$$

- $\pi(\theta)$ captures any prior information that is known about the parameters before calibration
- $\pi(d|\theta)$ is the likelihood that the parameters will be able to reproduce the data
- $\pi(d)$ is the model evidence
- $\pi(\theta|d)$ is the updated distribution of parameter values

Calibration of Model Parameters

Calibrate model parameters θ via Bayes' Rule

$$\underbrace{\pi(\theta|d)}_{\text{Posterior}} = \frac{\pi(d|\theta)\pi(\theta)}{\pi(d)}$$

- $\pi(\theta)$ captures any prior information that is known about the parameters before calibration
- $\pi(d|\theta)$ is the likelihood that the parameters will be able to reproduce the data
- $\pi(d)$ is the model evidence
- $\pi(\theta|d)$ is the updated distribution of parameter values

Calibrate Discrepancy Model

Propagate the MAP through the model and calculate the misfit:

$$d_\delta(x) = d(x) - m(\hat{\theta}, x)$$

Suppose we have a set of possible discrepancy models

$$\mathcal{M} = \{\delta_1, \delta_2, \ldots, \delta_n\}$$

Each δ_j has its own likelihood and parameters c_j

We can write Bayes' Rule in expanded form:

$$\pi(c_j|d_\delta) = \frac{\pi(d_\delta|c_j, \delta_j)\pi(c_j|\delta_j)}{\pi(d_\delta|\delta_j)}$$

Calibrate Discrepancy Model

Propagate the MAP through the model and calculate the misfit:

$$d_\delta(x) = d(x) - m(\hat{\theta}, x)$$

Suppose we have a set of possible discrepancy models

$$\mathcal{M} = \{\delta_1, \delta_2, \ldots, \delta_n\}$$

Each δ_j has its own likelihood and parameters c_j

We can write Bayes' Rule in expanded form:

$$\pi(c_j|d_\delta) = \frac{\pi(d_\delta|c_j, \delta_j)\pi(c_j|\delta_j)}{\pi(d_\delta|\delta_j)}$$

The evidence becomes the likelihood in another formulation of Bayes' Rule

Calibrate Discrepancy Model

$$\rho_j = \pi(\delta_j | d_\delta) = \frac{\pi(d_\delta | \delta_j)\pi(\delta_j)}{\pi(d_\delta)}$$

- $\pi(\delta_j)$ is the probability that model δ_j is the best model before calibration begins
- $\pi(d_\delta | \delta_j)$ is the evidence from the calibration of c_j
- $\pi(d_\delta)$ is a normalization constant
- ρ_j is the model plausibility

Calibrate Discrepancy Model

$$\rho_j = \pi(\delta_j | d_\delta) = \frac{\pi(d_\delta | \delta_j)\pi(\delta_j)}{\pi(d_\delta)}$$

- $\pi(\delta_j)$ is the probability that model δ_j is the best model before calibration begins
- $\pi(d_\delta | \delta_j)$ is the evidence from the calibration of c_j
- $\pi(d_\delta)$ is a normalization constant
- ρ_j is the model plausibility

Calibrate Discrepancy Model

$$\rho_j = \pi(\delta_j | d_\delta) = \frac{\pi(d_\delta | \delta_j)\pi(\delta_j)}{\pi(d_\delta)}$$

- $\pi(\delta_j)$ is the probability that model δ_j is the best model before calibration begins
- $\pi(d_\delta | \delta_j)$ is the evidence from the calibration of c_j
- $\pi(d_\delta)$ is a normalization constant
- ρ_j is the model plausibility

Calibrate Discrepancy Model

$$\rho_j = \pi(\delta_j | d_\delta) = \frac{\pi(d_\delta | \delta_j)\pi(\delta_j)}{\pi(d_\delta)}$$

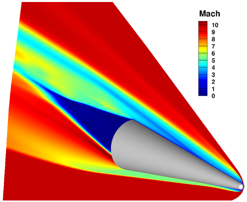
- $\pi(\delta_j)$ is the probability that model δ_j is the best model before calibration begins
- $\pi(d_\delta | \delta_j)$ is the evidence from the calibration of c_j
- $\pi(d_\delta)$ is a normalization constant
- ρ_j is the model plausibility

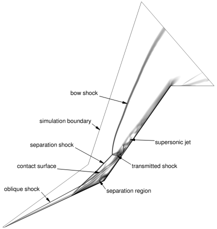
Calibrate Discrepancy Model

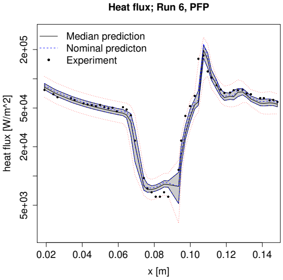
$$\rho_j = \pi(\delta_j | d_\delta) = \frac{\pi(d_\delta | \delta_j)\pi(\delta_j)}{\pi(d_\delta)}$$

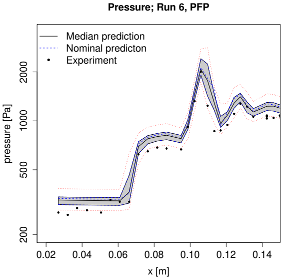
- $\pi(\delta_j)$ is the probability that model δ_j is the best model before calibration begins
- $\pi(d_\delta | \delta_j)$ is the evidence from the calibration of c_j
- $\pi(d_\delta)$ is a normalization constant
- ρ_j is the model plausibility

Application: Laminar Hypersonic Double-Cone Experiments





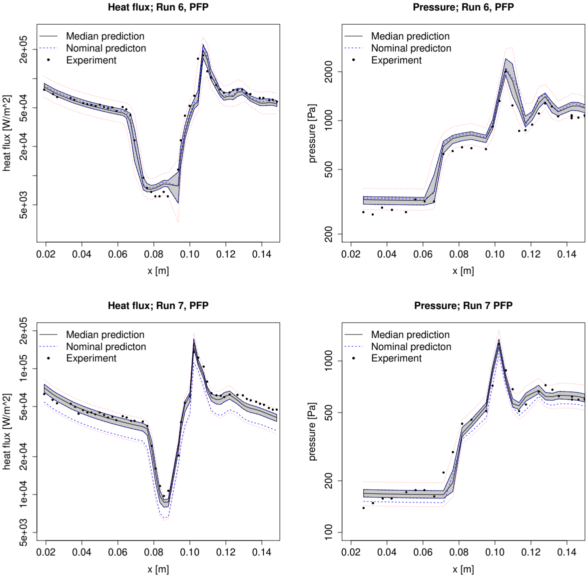




- d = LENS-I (Calspan-University at Buffalo Research Center)
- m = SPARC (Sandia National Labs)
- $\theta = \{\rho, U_\infty, T_\infty\}$
- Output QoI $= \{p, q, H_0, P_{pitot}\}$
- Previous calibration efforts revealed some discrepancy

Special thanks to Jaideep Ray and Sarah Kieweg. Ray, *et al.*, *AIAAJ* 2020

Calibration of Model Parameters



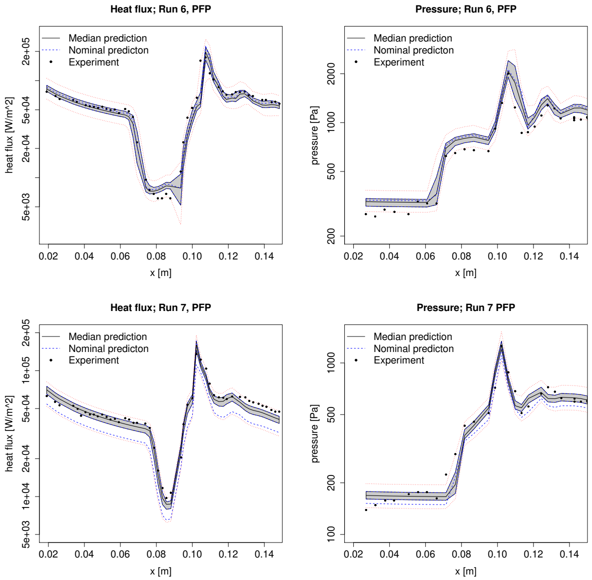
Constructed GP surrogate models for each QoI $\hat{m} = \{\hat{m}_p, \hat{m}_q, \hat{m}_{H_0}, \hat{m}_{P_{pitot}}\}$

- Built with $\sim$ 400 build points

Calibrated θ using Bayes' Rule

- $\pi(\theta) \sim$ uniform
- $\pi(d|\theta) \sim$ Gaussian
- $d = \{p, q, H_0, P_{pitot}\}$

Calibration of Model Parameters



Constructed GP surrogate models for each QoI $\hat{m} = \{\hat{m}_p, \hat{m}_q, \hat{m}_{H_0}, \hat{m}_{P_{pitot}}\}$

- Built with $\sim$ 400 build points

Calibrated θ using Bayes' Rule

- $\pi(\theta) \sim$ uniform
- $\pi(d|\theta) \sim$ Gaussian
- $d = \{p, q, H_0, P_{pitot}\}$

We used multiplicative discrepancy

$$d(x) = \hat{m}(\theta, x)\delta(x) + \varepsilon$$

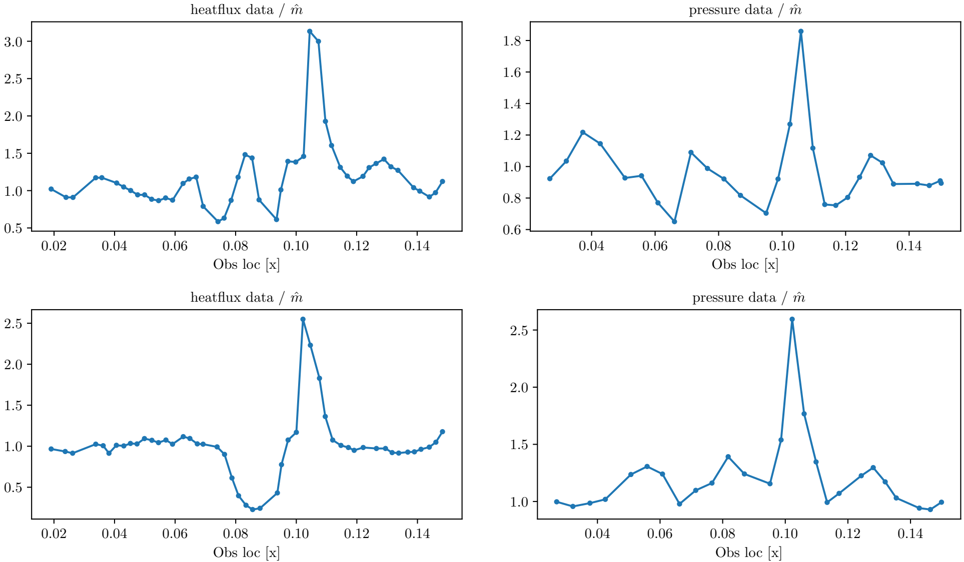
So, the misfit is calculated as

$$d_\delta(x) = \frac{d(x)}{\hat{m}(\hat{\theta}, x)}$$

Calibration of Discrepancy Model

Run 06

Run 07



Discrepancy Model Set

We defined the following set of possible discrepancy models

- Linear: $\delta(x, c) = c_0 x + c_1$
- Quadratic: $\delta(x, c) = (c_0 x + 0.5 - c_1)^2 + c_2$
- Logarithmic: $\delta(x, c) = c_0 \ln(c_1 x + 1 + c_2)$
- Exponential: $\delta(x, c) = \exp(c_0 x) - 1 + c_1$
- Gaussian: $\delta(x, c) = c_0 \exp\left(-\frac{1}{2}\frac{(x-c_1)^2}{c_2^2}\right) + c_3$
- Trigonometric: $\delta(x, c) = c_0 \sin(2\pi r) + c_3,\ r = \left[\frac{x-c_1}{c_2}\right]$

Discrepancy Model Set

We defined the following set of possible discrepancy models

- Linear: $\delta(x, c) = c_0 x + c_1$
- Quadratic: $\delta(x, c) = (c_0 x + 0.5 - c_1)^2 + c_2$
- Logarithmic: $\delta(x, c) = c_0 \ln(c_1 x + 1 + c_2)$
- Exponential: $\delta(x, c) = \exp(c_0 x) - 1 + c_1$
- Gaussian: $\delta(x, c) = c_0 \exp\left(-\frac{1}{2}\frac{(x-c_1)^2}{c_2^2}\right) + c_3$
- Trigonometric: $\delta(x, c) = c_0 \sin(2\pi r) + c_3,\ r = \left[\frac{x-c_1}{c_2}\right]$
- Gaussian + Trigonometric
- Sinc: $\delta(x, c) = c_0 \frac{\sin(2\pi r)}{2\pi r} + c_3,\ r = \left[\frac{x-c_1}{c_2}\right]$

Model log(evidences)

Functional Form	Heatflux		Pressure	
	Run 06	Run 07	Run 06	Run 07
Linear	-575	-623	-241	-212
Quadratic	-575	-581	-239	-211
Logarithmic	-568	-620	-236	-208
Exponential	-575	-623	-241	-212
Gaussian	-551	-608	-236	-166
Trigonometric	-567	**-552**	-218	-214
Trigonometric+Gaussian	-543	-609	-214	**-165**
Sinc	**-537**	-560	**-209**	-170

Model log(evidences)

Functional Form	Heatflux		Pressure	
	Run 06	Run 07	Run 06	Run 07
Linear	-575	-623	-241	-212
Quadratic	-575	-581	-239	-211
Logarithmic	-568	-620	-236	-208
Exponential	-575	-623	-241	-212
Gaussian	-551	-608	-236	-166
Trigonometric	-567	**-552**	-218	-214
Trigonometric+Gaussian	-543	-609	-214	**-165**
Sinc	**-537**	-560	**-209**	-170

Top Discrepancy Models - Heatflux

Run 06

Discrepancy

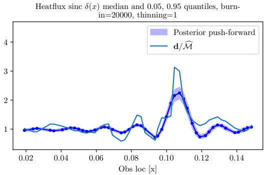
Data

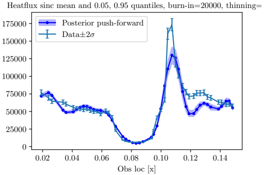
Run 07

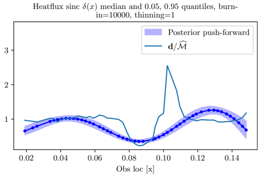
Discrepancy

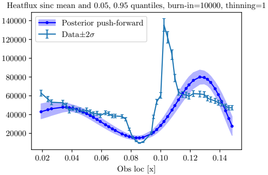
Data

Sinc

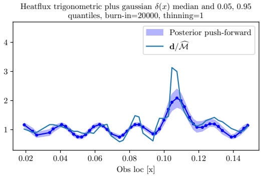


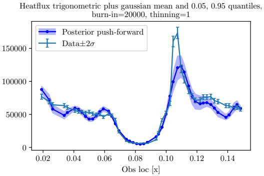


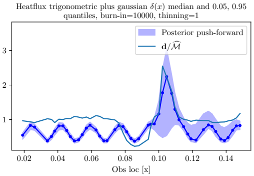


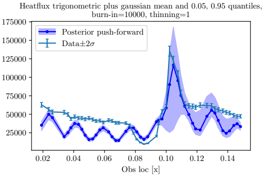


Trig + Gauss





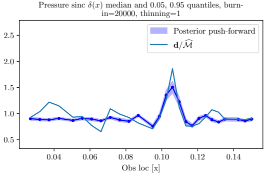


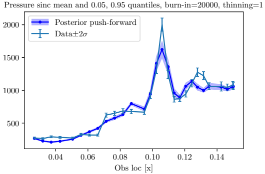


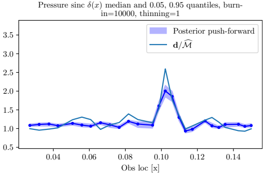
Top Discrepancy Models - Pressure

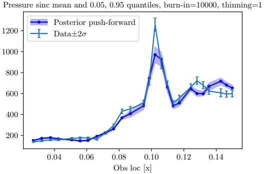
	Run 06		Run 07	
	Discrepancy	Data	Discrepancy	Data

Sinc

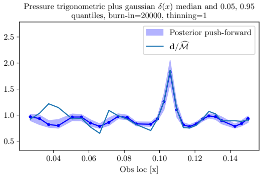


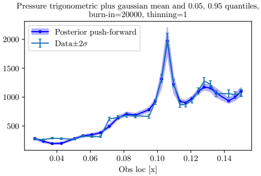


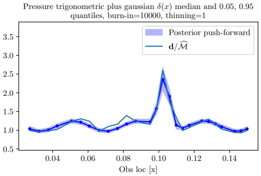


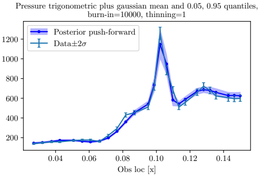


Trig + Gauss









Next step: Extrapolation

How can we extrapolate this information to new scenarios?

See poster by W. Lewis, et al. Multi-task Machine Learning for Fusion Simulations

Next step: Extrapolation

How can we extrapolate this information to new scenarios?

Multi-output Gaussian processes

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

Single-output

$$f_i(\mathbf{x}) \sim \mathcal{N}(\mu_i(\mathbf{x}), \boldsymbol{\Sigma}_i(\mathbf{x}, \mathbf{x}'))$$

$$\boldsymbol{\Sigma}_i \sim k_i(\mathbf{x}, \mathbf{x}')$$

$$\Rightarrow$$

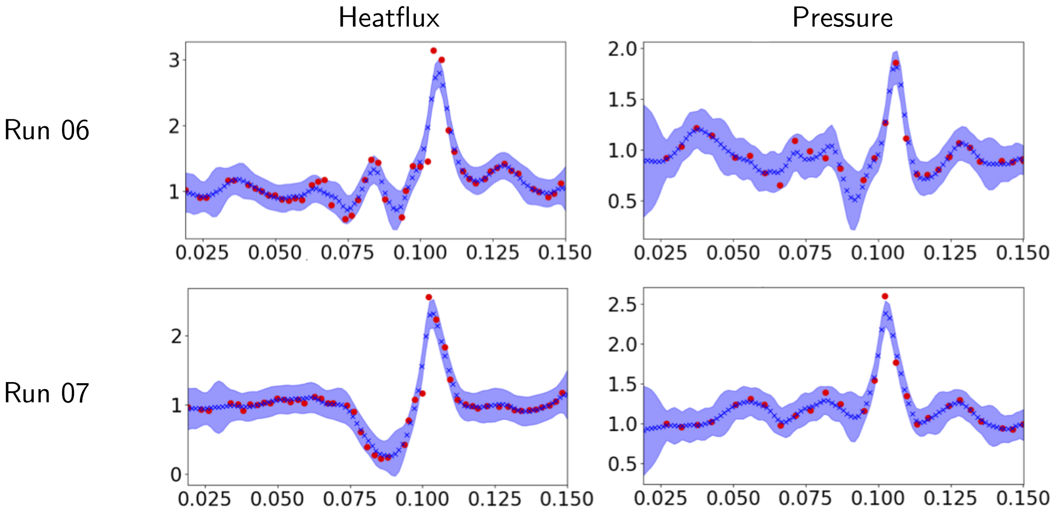
Multi-output

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

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

See poster by W. Lewis, et al. Multi-task Machine Learning for Fusion Simulations

Multi-output MFE



Summary

Combining model discrepancy with model selection yields a methodology for identifying missing physics

Successfully applied proposed approach to laminar hypersonic double cone experimental data

Next step: Extrapolate MFE to new experimental scenarios using MOGPs

Thank you

Questions?

kmaupin@sandia.gov

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

Two special cases

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

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

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