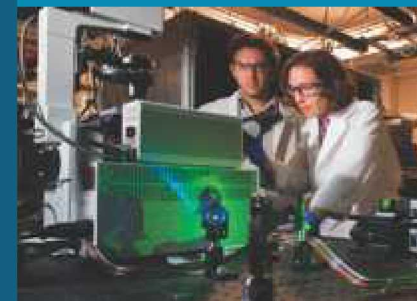
SAND2019-8178C xxC

Bayesian Calibration of Empirical Models Common in MELCOR and Other Nuclear Safety Codes



PRESENTED BY

Nathan W. Porter Postdoctoral Appointee

Vincent A. Mousseau Principle Member of the Technical Staff

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 1/2

- Empirical models used in reactor calculations are often based on data from decades ago.
- In many cases, these correlations are constructed using disparate data sets.
  - Different measurement devices/techniques
  - Various working fluids
  - Fluid property evaluations
  - Untreated physics/insufficient scaling
  - Human error
- The biases between these datasets are, in general, not treated during the formulation and calibration of the empirical model.

Phenomenon	Model	Year
Single phase Pressure drop	Colebrook Equation	1938
Two-phase pressure drop	Beattie	1982
Single phase wall heat transfer	Dittus-Boelter (McAdams)	1985 (1954)
Nucleate boiling heat transfer	Rohsenow	1951
Film boiling heat transfer	Bromley	1948
Critical heat flux	Zuber	1958
Film boiling	Zuber/Berenson	1960
Interfacial heat transfer	Lee & Ryley	1968
Terminal Taylor bubble velocity	Griffith	1961

Motivation 2/2

- Code uncertainty is underestimated when biases between datasets are not considered.
- When uncertainties are untreated, it is equivalent to assuming they are zero; in these cases, there is no motivation to reduce uncertainties.
- The purpose of this work is twofold:
  - Quantify and understand these biases/uncertainties, and
  - Motivate researchers to revisit some of these problems using modern hardware, techniques, and data acquisition systems.

Background and Theory

Calibration: Bayesian Methods

- In general, calibration is a statistical method to infer unknown parameter values/distributions by observing state variables and corresponding data (physical or computational experiments).
- Bayesian methods allow for the incorporation of prior information from previous experiments or expert knowledge.
- Solves Bayes' formula, which formulates the desired posterior distribution in terms of the prior distribution and likelihood function.

$$\pi(\theta|y) = \frac{\mathcal{L}(y|\theta)\pi_o(\theta)}{\int_\Theta \mathcal{L}(y|\theta)\pi_o(\theta)d\theta}$$

- We employ sampling methods because (1) the denominator is difficult or impossible to integrate and (2) the product of the likelihood and prior cannot be easily sampled.
- Delayed Rejection Adaptive Metropolis (DRAM) is one such sampling method.[1,2]

1. H. Haario, M. E. Saksman, and J. Tamminen, "An Adaptive Metropolis Algorithm," Bernoulli, 7(2), pp. 223-242 (2001) doi: 10.2307/3318737.
2. A. Mira, "On Metropolis-Hastings Algorithm with Delayed Rejection," Metron, 59(3), pp. 231-241 (2001).

Statistical Model

- For calibration, fixed effects statistical models are generally used, where the experimental data is equal to some model with zero-mean Gaussian measurement noise.

$$y = f(x, \theta) + \varepsilon$$

- This work employs mixed-effects statistical models, where each parameter is a combination of global and random effects.

$$y = f(x, \theta + \beta) + \varepsilon$$

- Frequentist solution methods generally solve this problem through minimization of an approximated likelihood function.[1]
- Bayesian Calibration can be used to obtain estimates of the desired posteriors, using conditional probabilities and likelihood function from the literature.[2]
- Here, a hierarchical Metropolis-within-DRAM method is used.[3,4]
- The DRAM step estimates global parameters, and Metropolis estimates random parameters.

1. J. C. Pinheiro and D. M. Bates, "Approximations to the Log-Likelihood Function in Nonlinear Mixed-Effects Model," *J Comp Graphical Stat,* **4**(1) (1995).
2. J. C. Wakefield, et al., "Bayesian Analysis of Linear and Nonlinear Populations Models Using the Gibbs Sampler," *J Royal Stat Soc*, **43**(1) (1994).
3. K. L. Schmidt, Uncertainty Quantification for Mixed-Effects Models with Applications in Nuclear Engineering, PhD thesis, North Carolina State University (2016).
4. M. Laine, MCMC Toolbox for Matlab (2017), helios.fmi.fi/~lainema/mcmc/.

Methodology

1. Gather experimental data from the literature.
   - Some will be related to original dataset
   - This set of data can be expanded in future work
2. Formulate the statistical model.
   - Use initial frequentist optimization
   - Determine which parameters are fixed/global/random via information criteria minimization
3. Calibrate the statistical model to the experimental data.
   - Use the hierarchical Metropolis-within-DRAM algorithm
   - Burn-in determined based on rule-of-thumb (10^5 iterations)[1]
4. Examine the results.
   - Compare initial correlation, frequentist optimization, and Bayesian calibration.
   - Propagate chain through original model and construct 95% predictive intervals

1. R. C. Smith. *Uncertainty Quantification, Theory, Implementation, and Applications*. SIAM (2014)

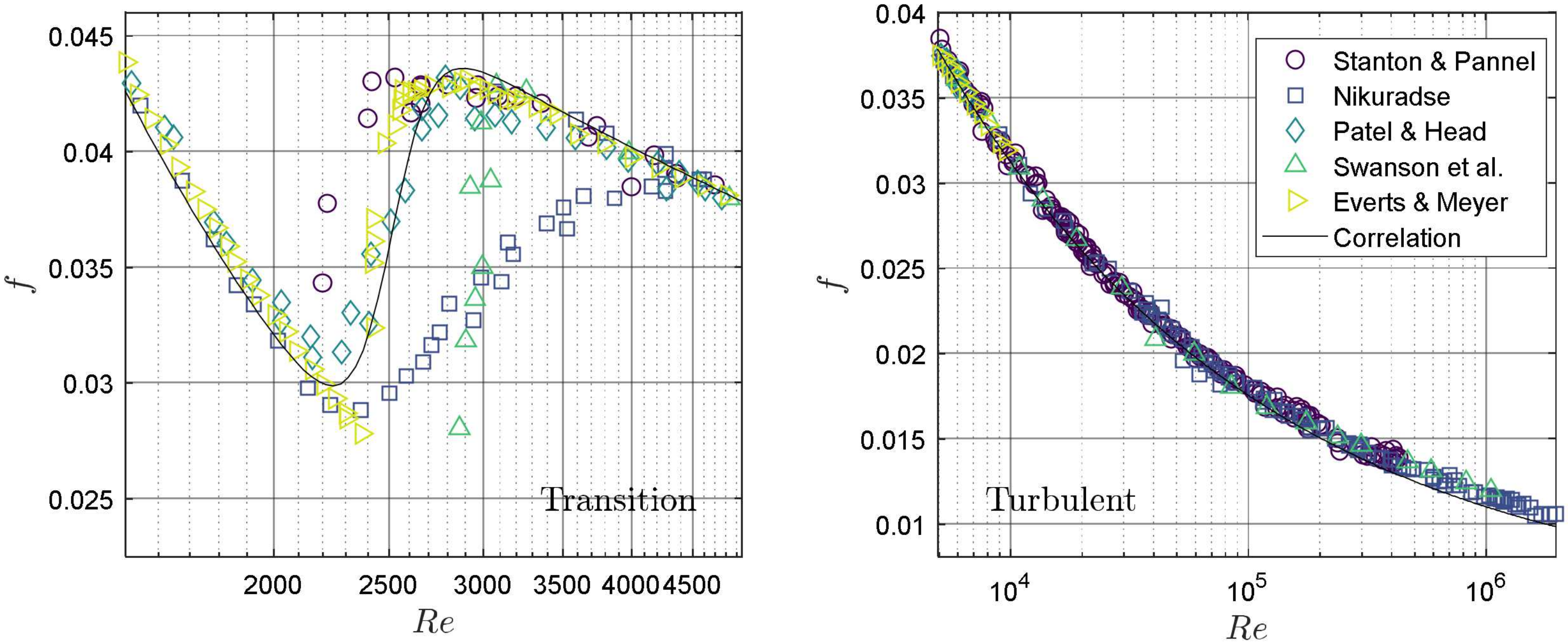
Results

Friction Factor 1/3

- Single phase friction in smooth tubes (Colebrook equation/Moody chart)
- Here, a new correlation is used which represents the transition region using a logistic function.

$$f = (1 - S) f_{laminar} + S f_{turbulent}$$

$$f = \left[1 - \frac{1}{1 + e^{-q_1(Re - q_2)}}\right] \frac{64}{Re} + \frac{q_3 + q_4 Re^{-q_5}}{1 + e^{-q_1(Re - q_2)}}$$



Friction Factor 2/3

- Through minimization of the information criteria, only q_1 and q_2 require random effects.

$$f = \left[1 - \frac{1}{1 + e^{-(\theta_1+\beta_{1l})\,[Re-(\theta_2+\beta_{2l})]}}\right]\frac{64}{Re} + \frac{\theta_3 + \theta_4 Re^{-\theta_5}}{1 + e^{-(\theta_1+\beta_{1l})\,[Re-(\theta_2+\beta_{2l})]}}$$

Global	Random	AIC	BIC
	q_1, q_2, q_3, q_4, q_5	-5297.7	-5302.0
q_5	q_1, q_2, q_3, q_4	-5299.7	-5303.6
q_4, q_5	q_1, q_2, q_3	-5301.7	-5305.2
q_3, q_4, q_5	q_1, q_2	-5303.7	-5306.8
q_2, q_3, q_4, q_5	q_1	-4770.1	-4772.8
q_1, q_2, q_3, q_4, q_5		-4515.4	-4517.8

Friction Factor 2/3

- Through minimization of the information criteria, only q_1 and q_2 require random effects.

$$f = \left[1 - \frac{1}{1 + e^{-(\theta_1+\beta_{1l})\,[Re-(\theta_2+\beta_{2l})]}}\right]\frac{64}{Re} + \frac{\theta_3 + \theta_4 Re^{-\theta_5}}{1 + e^{-(\theta_1+\beta_{1l})\,[Re-(\theta_2+\beta_{2l})]}}$$

Global	Random	AIC	BIC
	q_1, q_2, q_3, q_4, q_5	-5297.7	-5302.0
q_5	q_1, q_2, q_3, q_4	-5299.7	-5303.6
q_4, q_5	q_1, q_2, q_3	-5301.7	-5305.2
q_3, q_4, q_5	q_1, q_2	-5303.7	-5306.8
q_2, q_3, q_4, q_5	q_1	-4770.1	-4772.8
q_1, q_2, q_3, q_4, q_5		-4515.4	-4517.8
q_3, q_5	q_1, q_2, q_4	-5301.7	-5305.2
$q_3, q_4,$	q_1, q_2, q_5	-5301.7	-5305.2

Friction Factor 2/3

Global	Random	AIC	BIC
	q_1, q_2, q_3, q_4, q_5	-5297.7	-5302.0
q_5	q_1, q_2, q_3, q_4	-5299.7	-5303.6
q_4, q_5	q_1, q_2, q_3	-5301.7	-5305.2
q_3, q_4, q_5	q_1, q_2	-5303.7	-5306.8
q_2, q_3, q_4, q_5	q_1	-4770.1	-4772.8
q_1, q_2, q_3, q_4, q_5		-4515.4	-4517.8
q_3, q_5	q_1, q_2, q_4	-5301.7	-5305.2
$q_3, q_4,$	q_1, q_2, q_5	-5301.7	-5305.2

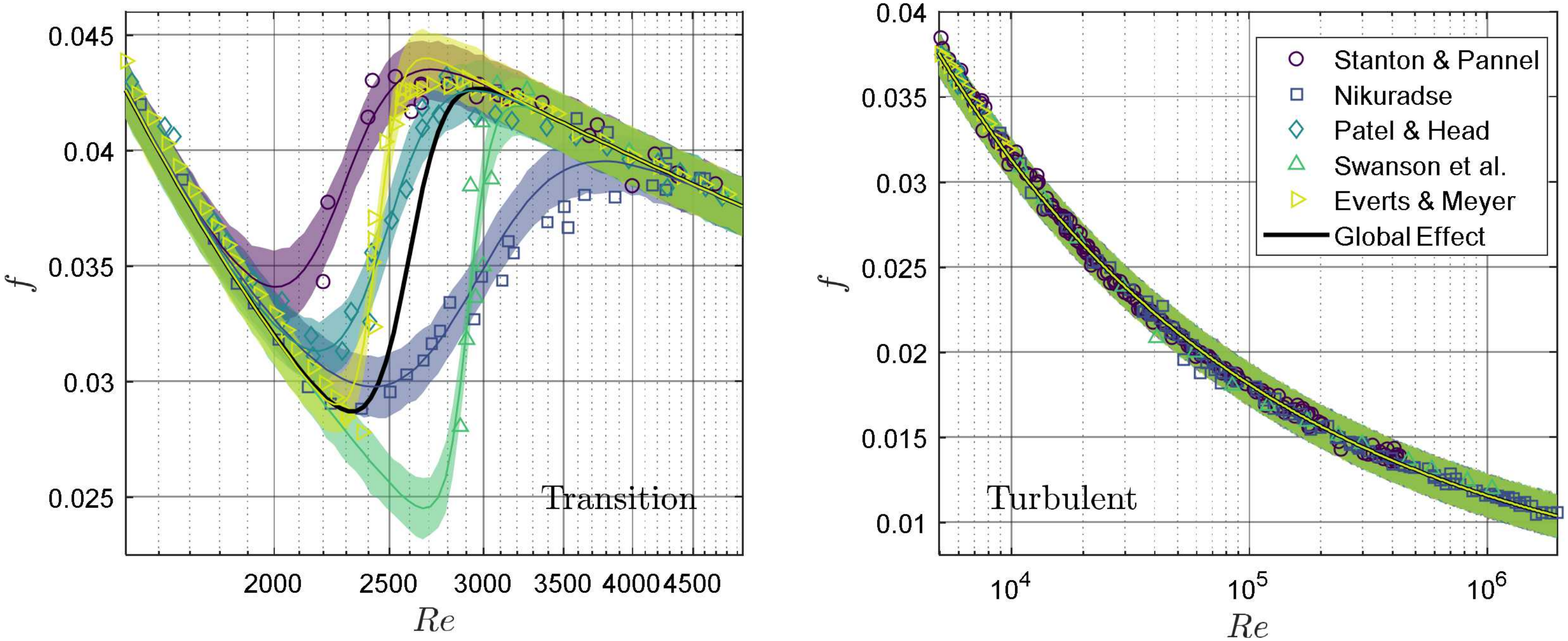
- Through minimization of the information criteria, only q_1 and q_2 require random effects.

$$f = \left[1 - \frac{1}{1 + e^{-(\theta_1 + \beta_{1l})\,[Re - (\theta_2 + \beta_{2l})]}}\right]\frac{64}{Re} + \frac{\theta_3 + \theta_4 Re^{-\theta_5}}{1 + e^{-(\theta_1 + \beta_{1l})\,[Re - (\theta_2 + \beta_{2l})]}}$$

		Frequentist Optimization					**Bayesian Calibration (mean results)**				
		Stanton	Nikuradse	Patel	Swanson	Everts	Stanton	Nikuradse	Patel	Swanson	Everts
θ_1	0.01	0.01022					0.01042				
β_{1l}		-0.00269	-0.00667	-0.00289	0.00480	0.00745	-0.00285	-0.00685	-0.00299	0.00472	0.00776
θ_2	2500	2582					2584				
β_{2l}		-331.5	272.4	-124.3	325.8	-142.4	-334.7	270.7	-126.1	324.3	-143.9
θ_3	0.005	0.005169					0.005156				
θ_4	0.5	0.4354					0.4336				
θ_5	0.32	0.3054					0.3048				
σ^2		5.996e-7					5.940e-7				
Σ		$\begin{bmatrix} 2.844e-5 & 0 \\ 0 & 6.552e4 \end{bmatrix}$					$\begin{bmatrix} 4.663e-5 & 0 \\ 0 & 1.017e4 \end{bmatrix}$				

Friction Factor 3/3

- Through propagation of the Bayesian chains through the original model, uncertainty bounds can be found for each individual experiment.
- Here, the 95% prediction intervals are shown, and they capture the appropriate percentage of the data.
- Only the parameters which determine the transition region have random effects, so the laminar and turbulent regions do not have laboratory-dependent uncertainty.

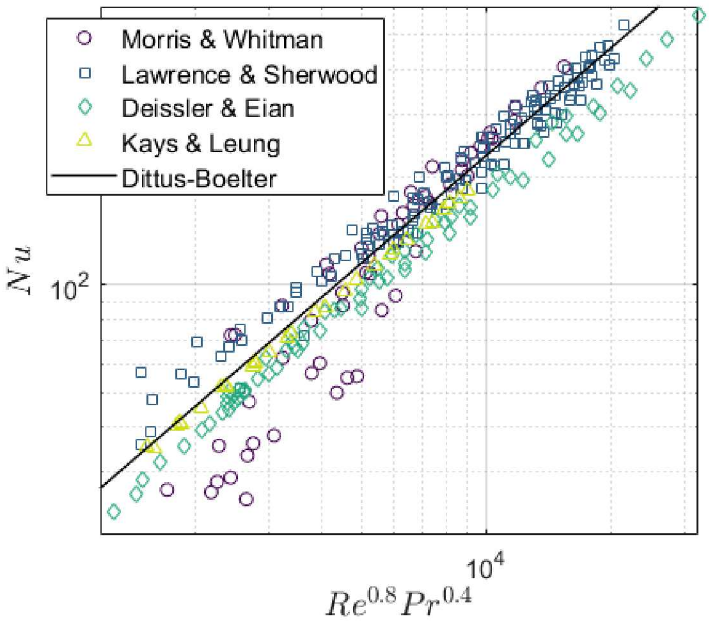


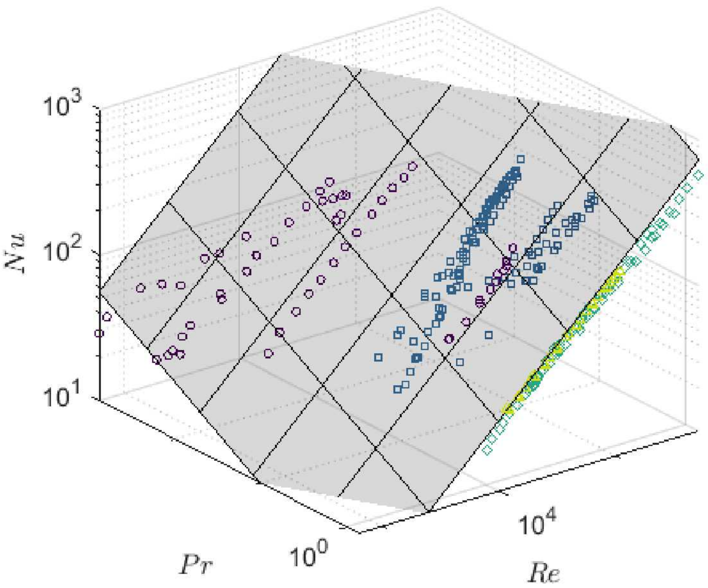
Heat Transfer Coefficient 1/3

- Single phase forced convection heat transfer in smooth tubes
- Represented by the Dittus-Bolter equation, which was created for analysis of automobile radiators

$$Nu = q_1 Re^{q_2} Pr^{q_3}$$

- Data is shown in both two and three dimensions





Heat Transfer Coefficient 2/3

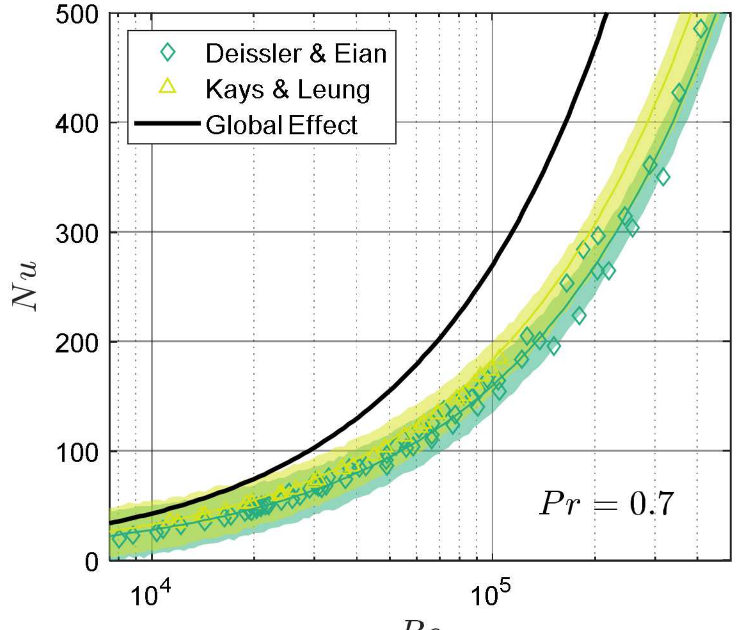
- Through minimizing AIC and BIC, all three parameters require random effects to fit the data.

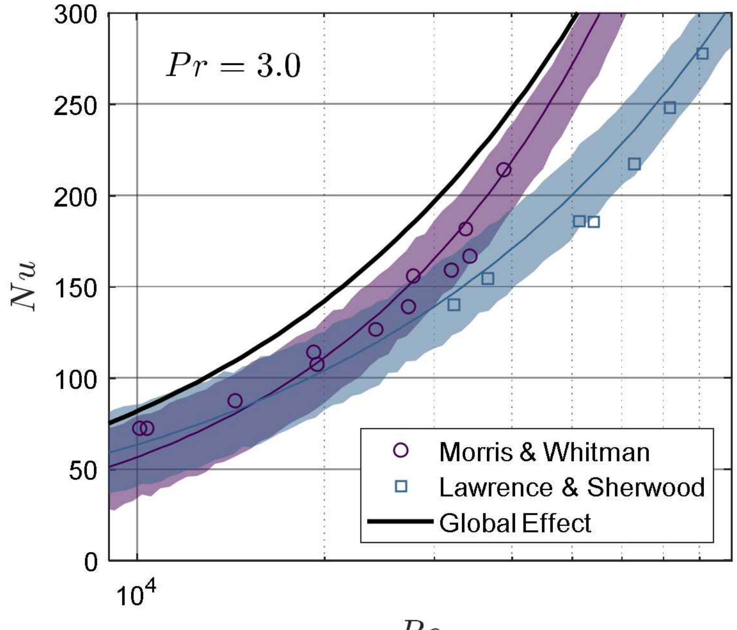
$$Nu = (\theta_1 + \beta_{1l})Re^{(\theta_2+\beta_{2l})}Pr^{(\theta_3+\beta_{3l})}$$

		Frequentist Optimization				Bayesian Calibration (mean results)			
		Morris	Lawrence	Deissler	Kays	Morris	Lawrence	Deissler	Kays
θ_1	0.023	0.0285				0.03234			
β_{1l}		-0.0244	0.0211	8.58e-4	0.00247	-0.0277	0.0203	-7.89e-4	0.00770
θ_2	0.8	0.8064				0.7978			
β_{2l}		0.176	-0.0877	-0.0470	-0.0410	0.176	-0.0842	-0.0428	-0.0524
θ_3	0.4	0.4431				0.4433			
β_{3l}		-0.0326	0.0327	-3.07e-5	-1.06e-3	-0.0379	0.0315	0.0191	-0.0128
σ^2		183.3				184.07			
Σ		$\begin{bmatrix} 3.065e-4 & 0 & 0 \\ 0 & 1.105e-2 & 0 \\ 0 & 0 & 1.187e-3 \end{bmatrix}$				$\begin{bmatrix} 5.570e-4 & 0 & 0 \\ 0 & 1.827e-2 & 0 \\ 0 & 0 & 3.469e-3 \end{bmatrix}$			

Heat Transfer Coefficient 3/3

- Again, the chains are propagated through the original model.
- It is difficult to plot three dimensional prediction intervals, so they are plotted for fixed Prandtl number.
- The appropriate fraction of data is within the prediction intervals.
- The global effect, which only uses the fixed parameters, is significantly different than any of the laboratory results.



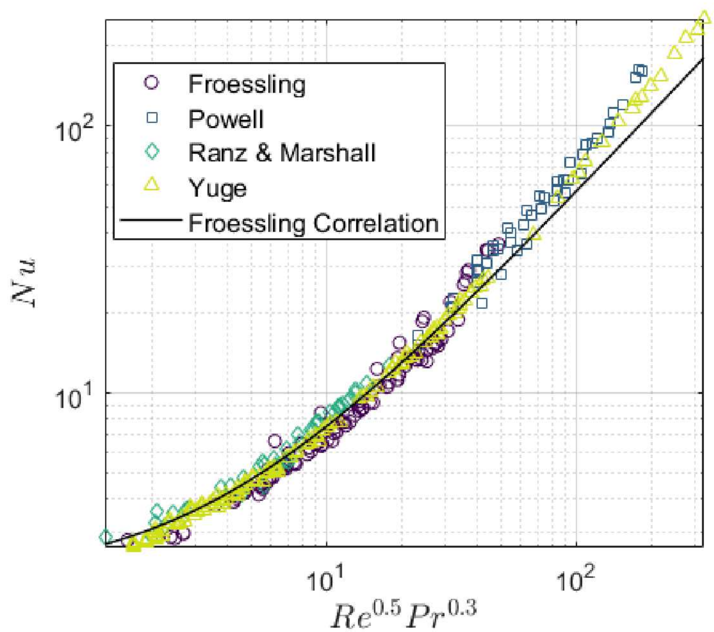


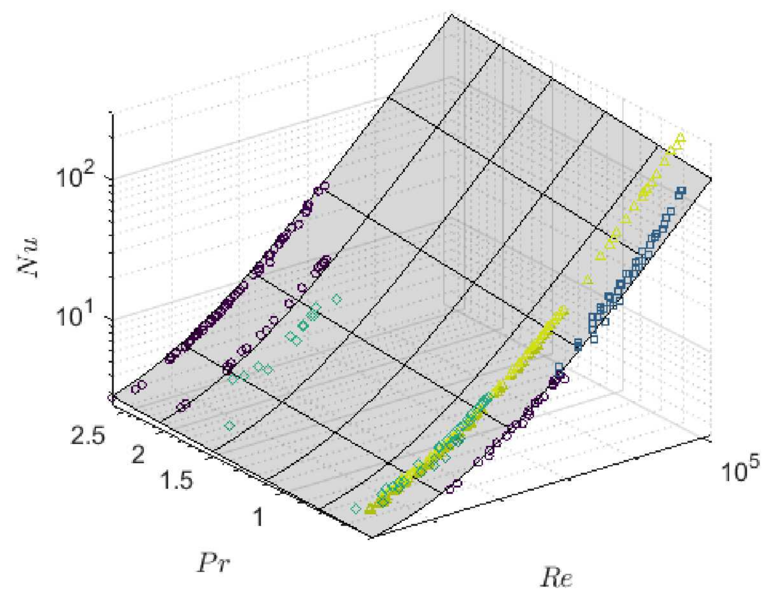
Mass Transfer Coefficient 1/3

- Mass or heat transfer from a solid sphere to the surrounding medium.
- Used to approximate interfacial transfer for bubbles/droplets.

$$Nu = 2.0 + \theta_1 Re^{\theta_2} Pr^{\theta_3}$$

- Originally proposed by Froessling (1938), but θ_1 has been adjusted by various authors; most recently, Ranz & Marshall (1952) and Lee & Ryley (1968).





Mass Transfer Coefficient 2/3

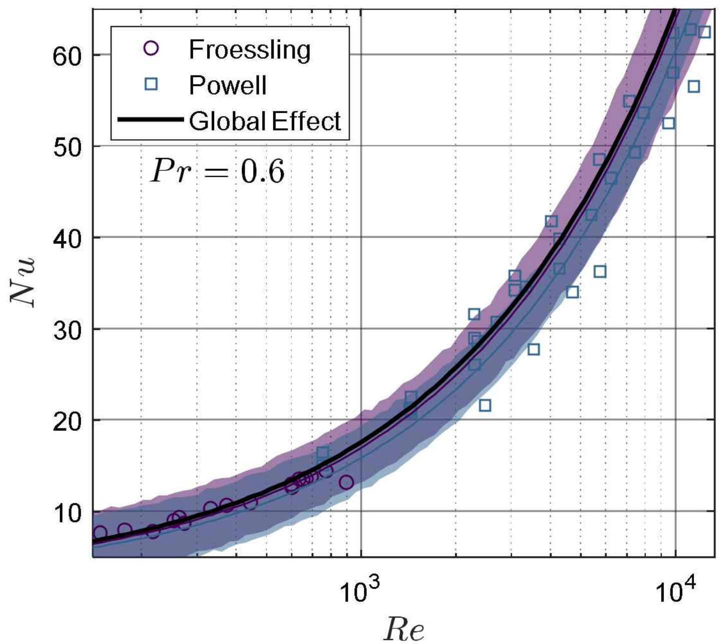
- The statistical model is formulated such that q_1 and q_2 have random effects.

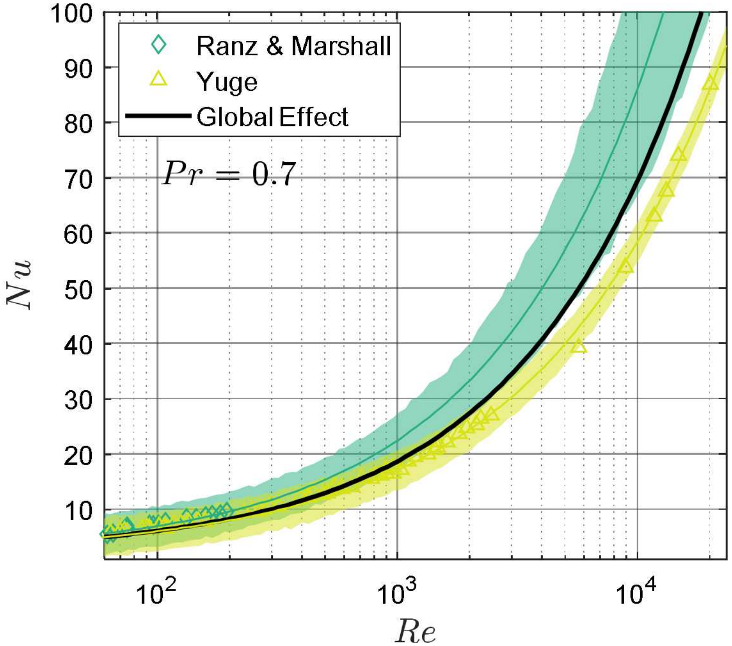
$$Nu = 2 + (\theta_1 + \beta_{1l})Re^{(\theta_2+\beta_{2l})}Pr^{\theta_3}$$

		Frequentist Optimization				**Bayesian Calibration (mean results)**			
		Froessling	Powell	Ranz	Yuge	Froessling	Powell	Ranz	Yuge
θ_1	0.55	0.2920				0.2957			
β_{1l}		-0.0275	-0.0625	0.0442	0.0458	-0.0310	-0.0653	0.0510	0.0439
θ_2	0.5	0.6080				0.6066			
β_{2l}		0.00853	0.0176	0.00930	-0.0355	0.0101	0.0189	0.00659	-0.0344
θ_3	0.33	0.4414				0.4436			
σ^2		4.621				4.672			
Σ		$\begin{bmatrix} 2.705e-3 & 0 \\ 0 & 6.002e-4 \end{bmatrix}$				$\begin{bmatrix} 5.972e-3 & 0 \\ 0 & 1.246e-3 \end{bmatrix}$			

Mass Transfer Coefficient 3/3

- Bayesian chains are propagated through the original model.
- Observational error should be laboratory-dependent, ε_l.
  - For Froessling and Ranz, all data is captured in the interval (error overestimated)
  - For Powell, half the data is captured (error underestimated)





Conclusion

- Some simple empirical relations have been examined using the legacy data from which they were derived (in some cases, this data is supplemented with other sources).
- Many possible sources of experimental bias, some hypotheses are supported by mixed effects results.
  - Transition to turbulence can be impacted by entry geometry, development, pump vibrations, and even working fluid.
  - Variety of working fluids in heat transfer experiments.
  - Choice of diffusion coefficients in the mass transfer experiments, which were very uncertain at the time.
- Biases between experiments have been quantified, which allows an accurate quantification of experimental uncertainty.
- Through reexamination and repeated experiments, it may be possible to reduce these uncertainties.

Future Work

- There are some possible improvements to the calibration method.
  - Random effects drawn from a non-normal distribution (e.g., Johnson distribution)
  - Dependent hyperparameters (Σ is currently diagonal).[1]
  - Nonlinear or laboratory-dependent observational error, ε.
  - Incorporation of more sophisticated parameter-selection algorithm.[1]
  - Use of empirical convergence criterion. [2]
- Incorporation of more recent data, which often includes more physics (pipe roughness, natural circulation, geometry, etc.).
- Application to other nuclear correlations used in nuclear codes, with specific focus on severe accident analysis.
- Use of results to quantify and exclude "poor" data.

1. K. L. Schmidt, Uncertainty Quantification for Mixed-Effects Models with Applications in Nuclear Engineering, PhD thesis, North Carolina State University (2016).
2. S. Brooks and A. Gelman, "General Methods for Monitoring Convergence of Iterative Simulations," *J Comput Graph Stat*, **7** (1998).
3. J. C. Wakefield, et al., "Bayesian Analysis of Linear and Nonlinear Populations Models Using the Gibbs Sampler," *J Royal Stat Soc*, **43**(1) (1994).

- Thanks to the following individuals for their assistance with mixed-effects Bayesian Calibration:

  North Carolina State University Paul R. Miles, Kathleen L. Schmidt, Ralph Smith

AIC and BIC

- Penalized-likelihood criteria that are often used to choose best predictor during regression analysis. Model with minimized value represents the best fit.
- Measure of fit + penalty for complexity
- Akaike information criterion (AIC)

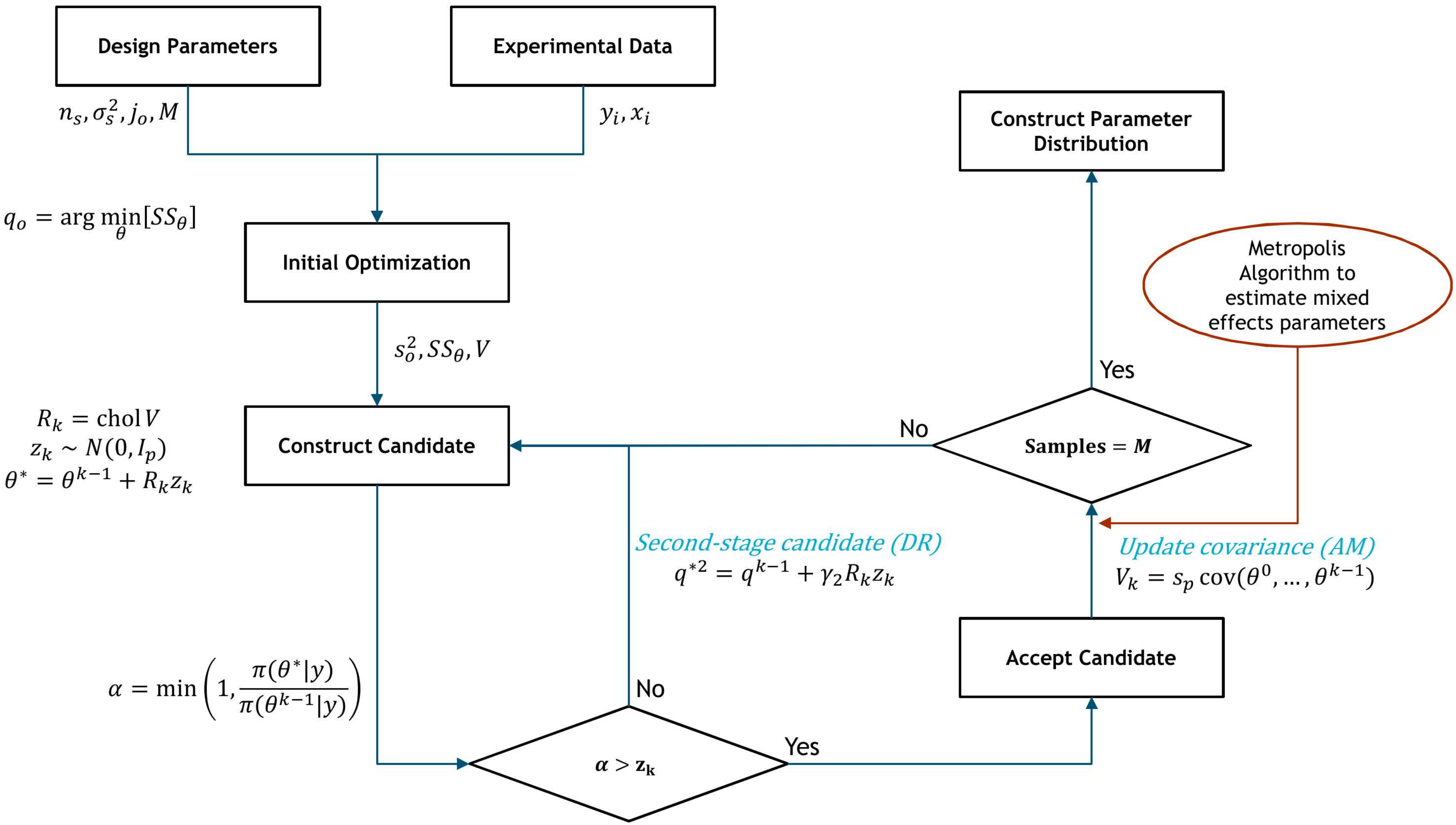
$$AIC = -2 \ln \mathcal{L} + 2p$$

- Bayesian information criterion (BIC)

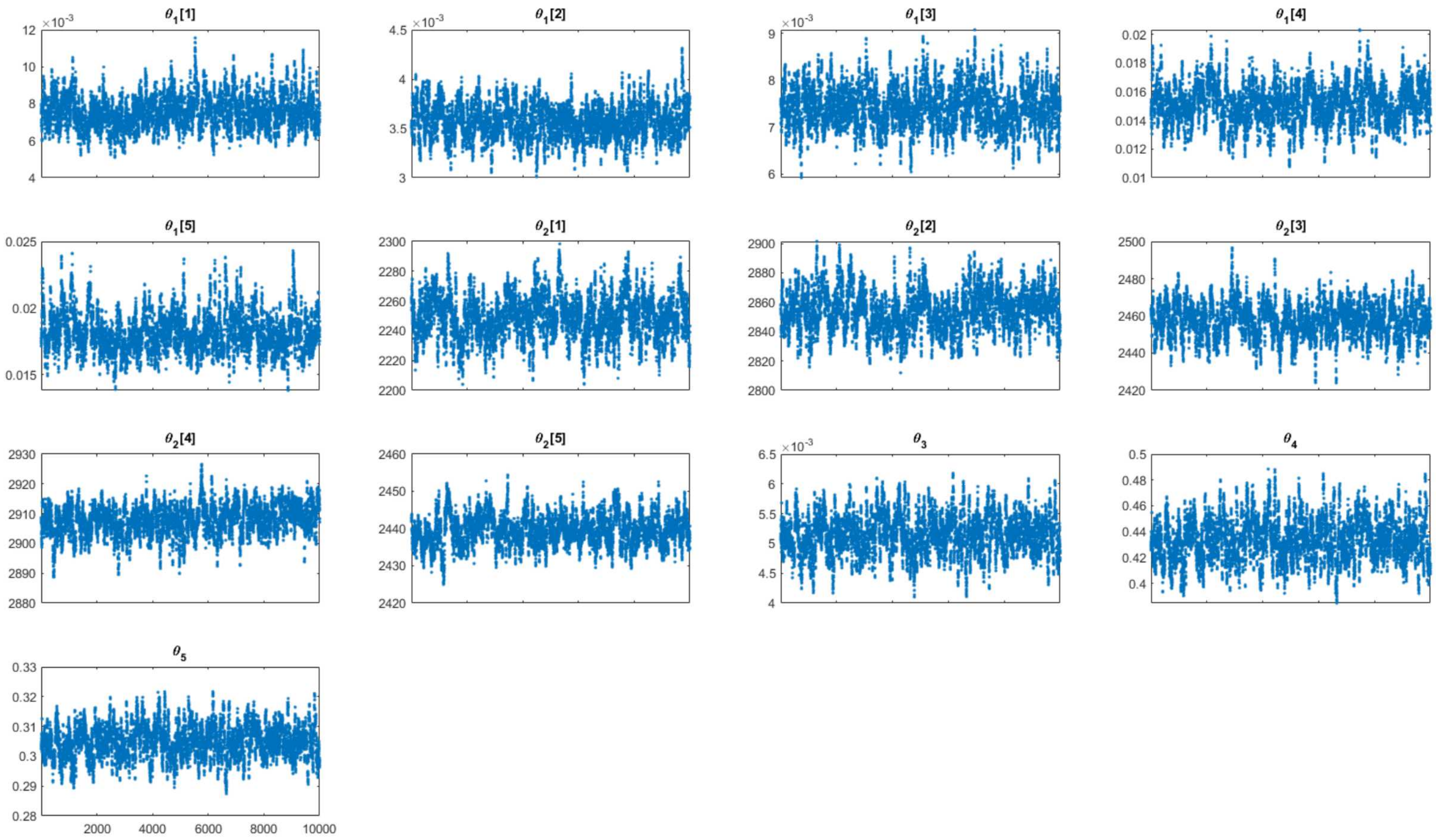
$$BIC = -2 \ln \mathcal{L} + p \ln N$$

$\mathcal{L}$: Likelihood
p: number of parameters
N: number of data points

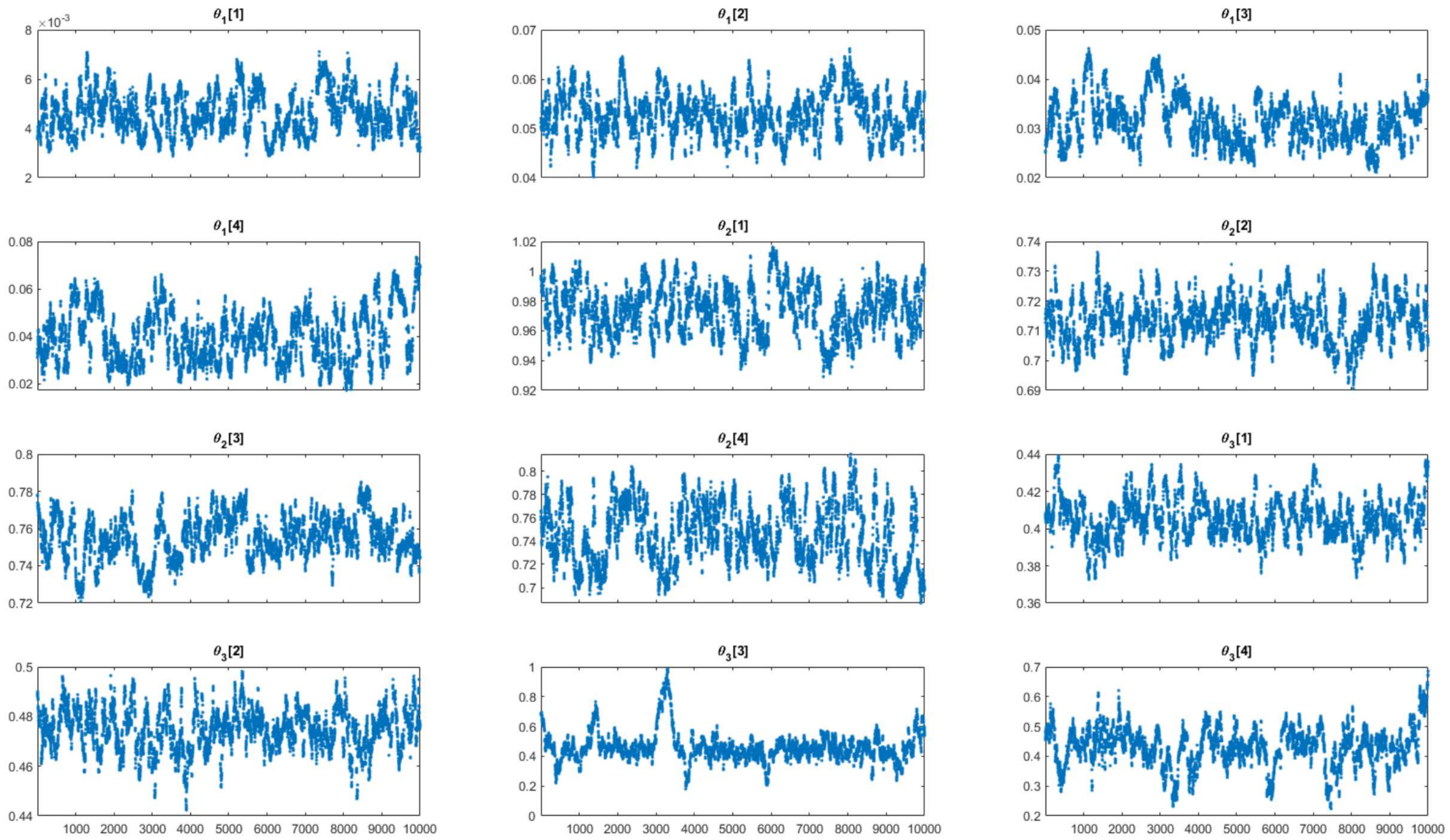
Metropolis-within-DRAM Algorithm



Friction Factor Chains



Heat Transfer Coefficient Chains



Mass Transfer