Estimating Computational Biases for Criticality Safety Applications with Few Neutronically Similar Benchmarks

Dr. Christopher M. Perfetti
Dr. Bradley T. Rearden
Dr. William J. Marshall
Radiation Transport Group
Reactor and Nuclear Systems Division
Oak Ridge National Laboratory

ORNL is managed by UT-Battelle
for the US Department of Energy

OAK RIDGE
National Laboratory

Introduction

- Criticality safety analyses rely on the availability of relevant benchmark experiments to determine justifiable margins of subcriticality.
- Validation efforts seek use benchmark experiments to estimate the computational bias in the predicted eigenvalue for applications.

$$Bias = \beta = k_{calc} - k_{exp}$$

Introduction

- This study is meant to compare the predictive capabilities of criticality safety validation approaches.
- This blind benchmark study applies predictive capabilities to low-moderated MOX powder experiments with few representative experiments.
  - This study was also performed using 10 cases with known biases.

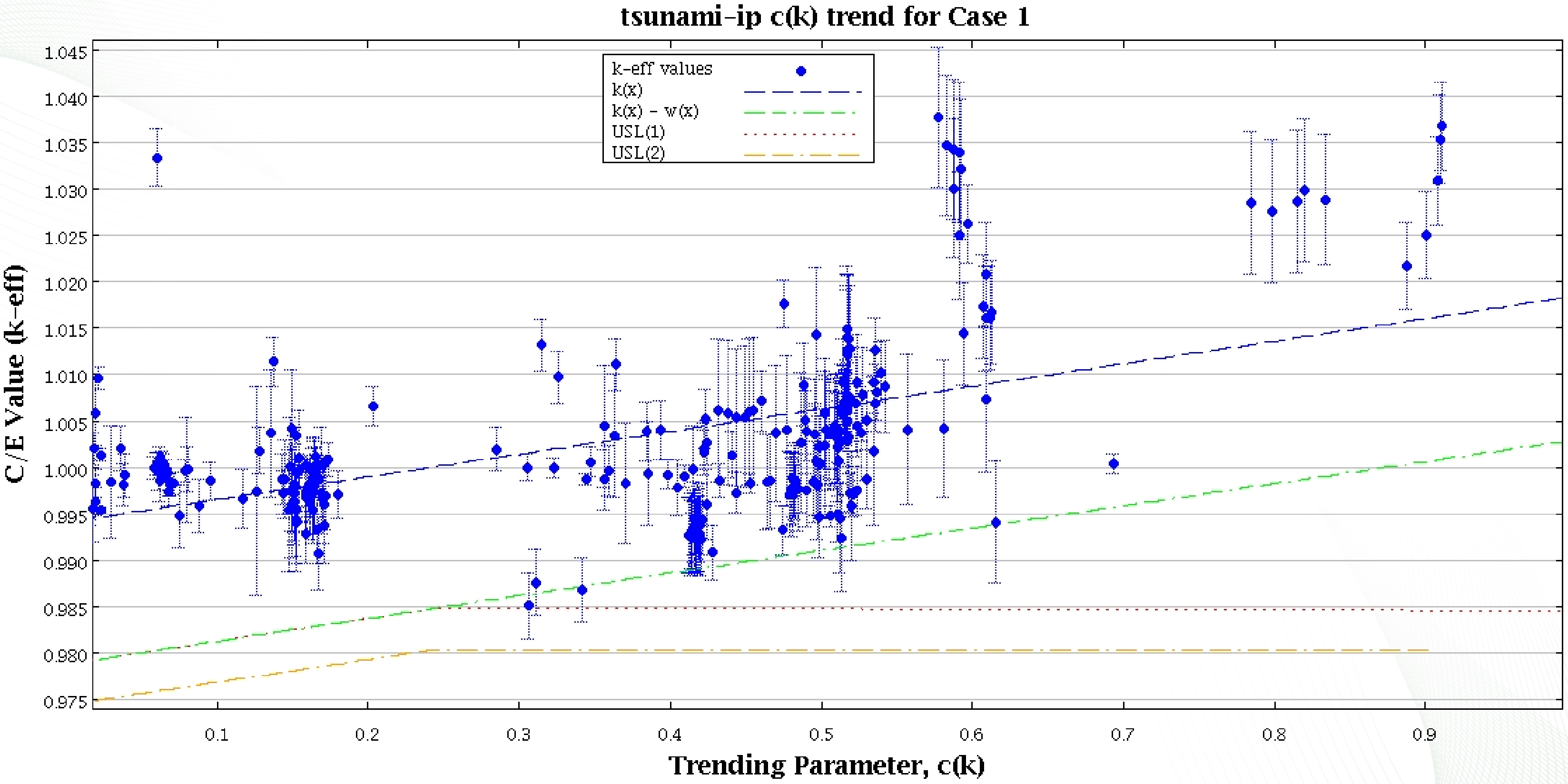
Introduction

- Our study compares three bias estimation methodologies:
  - Trending Analysis (USLSTATS, ORNL)
  - Non-parametric Methods (Whisper, LANL)
  - Experimental Data Assimilation (TSURFER, ORNL)

Trending Analysis – USLSTATS

- USLSTATS is a generic parameter trending analysis tool from ORNL.
- USLSTATS provides several predictive confidence parameters, including:
  1. Expected application bias
  2. Confidence band with administrative margin (USL1)
     - Does not give credit for positive biases.
  3. Single-sided, uniform-width confidence interval (USL2)
- The administrative margin was set to zero for this exercise for all methods.

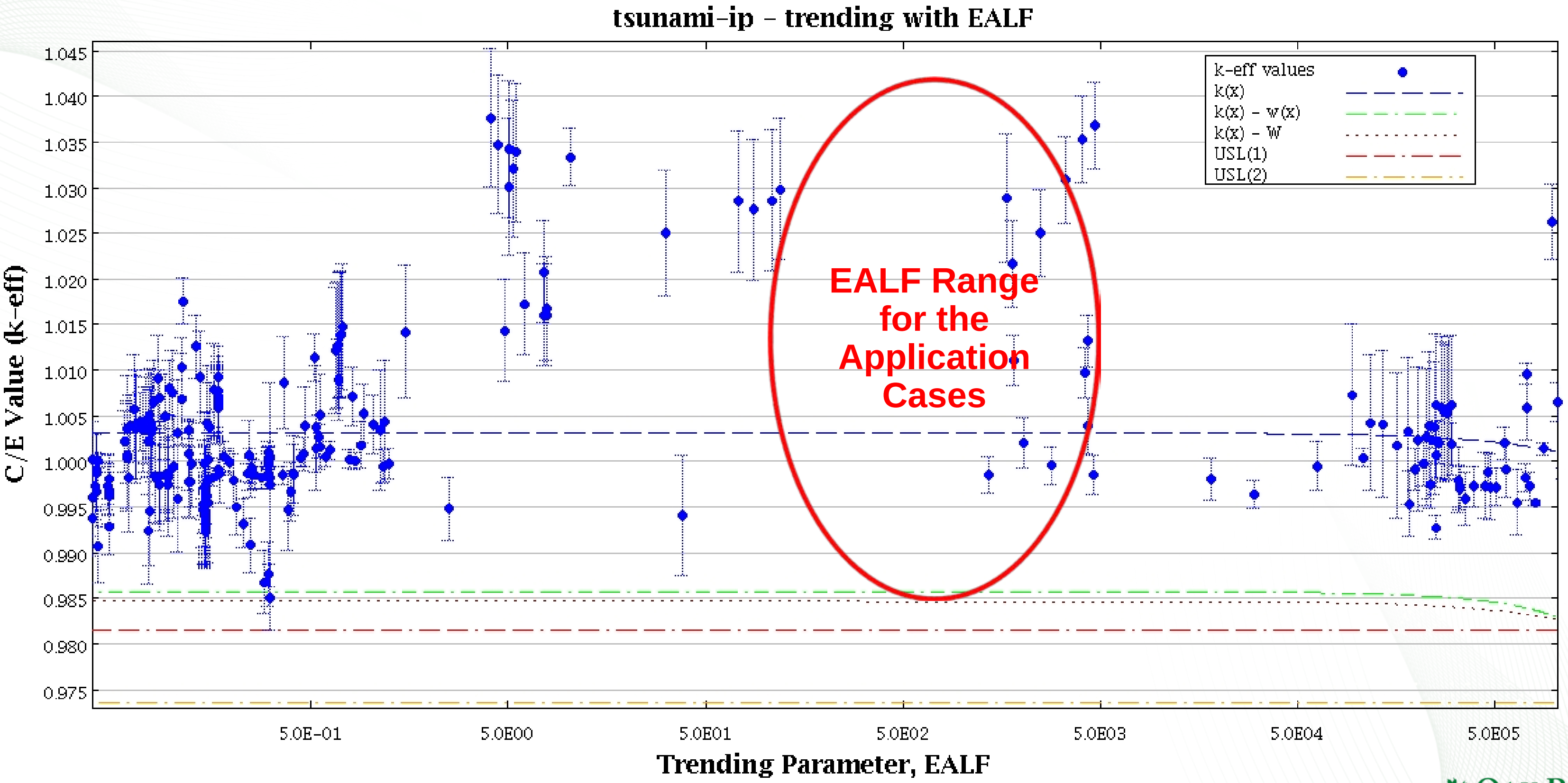
Trending Analysis – Sample USLSTATS Output



Trending Analysis – USLSTATS

- Trending parameters examined in this study included:
  1. The coefficient of similarity, c(k) or c_k
  2. The Energy of the Average Lethargy of Fission (EALF)
- Since the UACSA Phase V exercise is a blind benchmark study, we don’t know for sure what the correct answer is.
  – Cases with known biases were examined to explore the accuracy of the bias estimation methods.
  – For the unknown bias cases, emphasis will be placed on methods that produce consistent results.

Trending on EALF



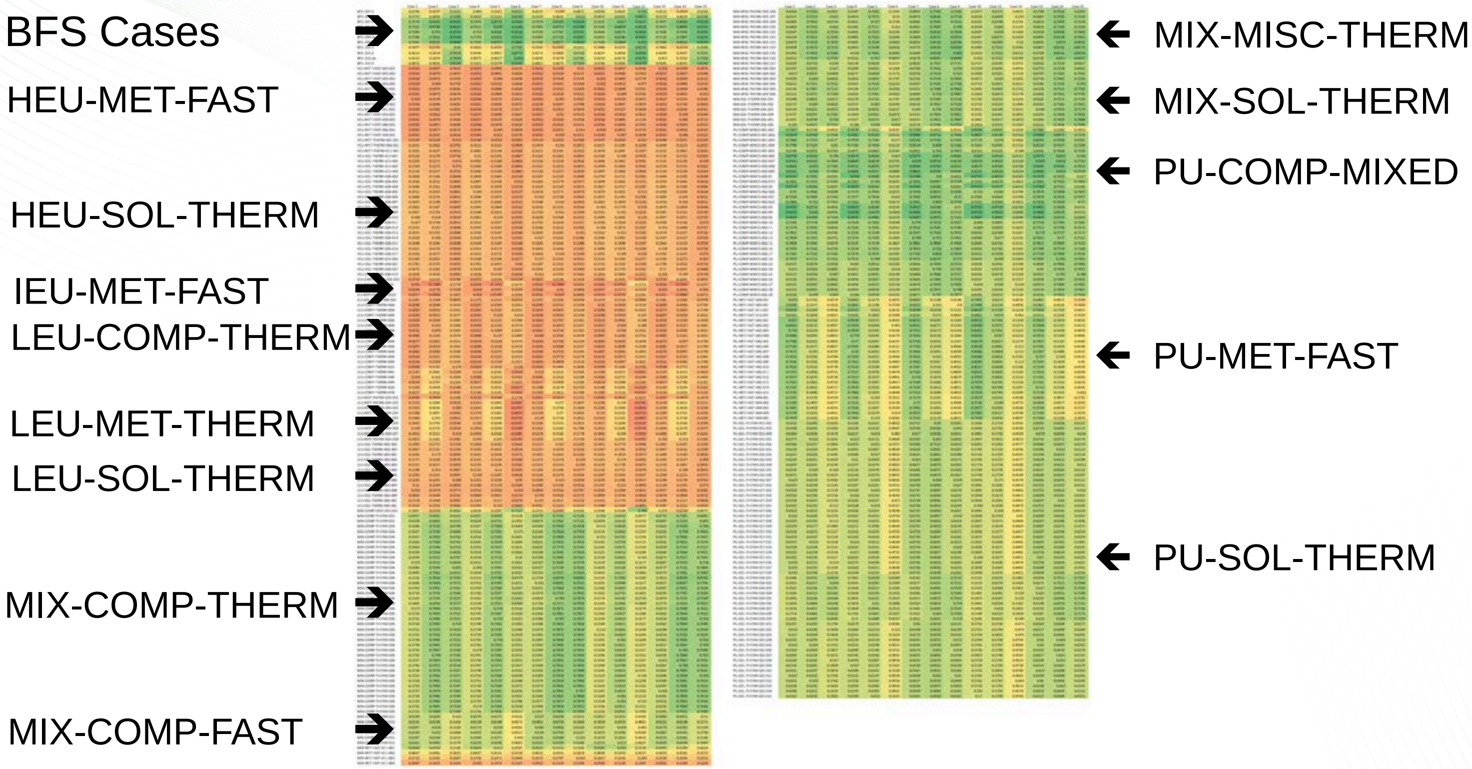
Benchmark Similarity Assessment

- The similarity coefficient, *c(k)* or c_k, describes the amount of nuclear data-induced uncertainty that is shared by two systems.

$$\underbrace{S_{R_1,\Sigma_x}}_{\left(\frac{\delta R/R}{\delta\Sigma/\Sigma}\right)} \cdot \underbrace{Cov_{\Sigma_x,\Sigma_y}}_{(\Delta\Sigma/\Sigma)^2} \cdot \underbrace{S^T_{R_2,\Sigma_y}}_{\left(\frac{\delta R/R}{\delta\Sigma/\Sigma}\right)} = \underbrace{\sigma^2_{R_1,R_2}}_{(\Delta R/R)^2}$$

$$c_k = \frac{\sigma^2_{R_1,R_2}}{\sigma_{R_1}\sigma_{R_2}}$$

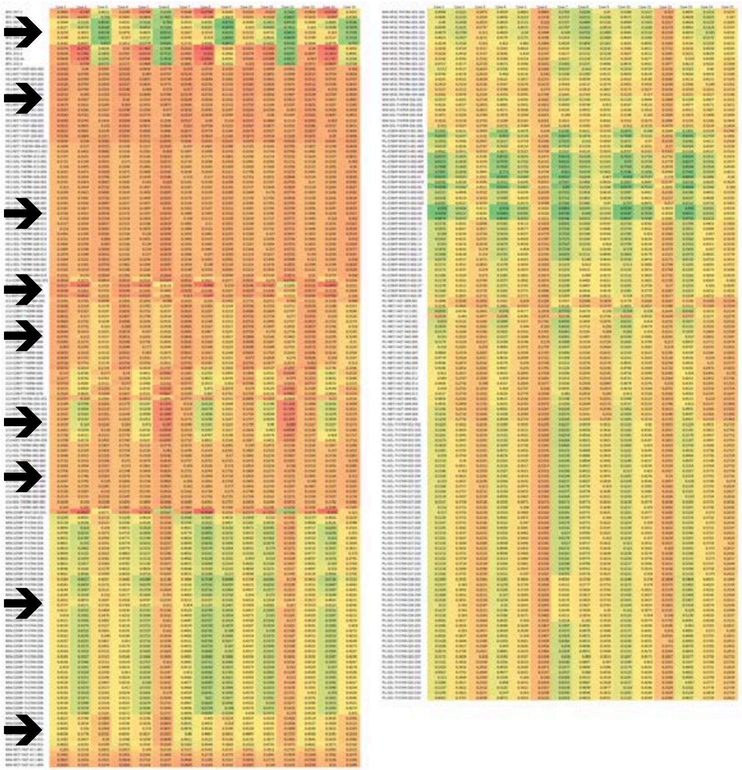
Benchmark Experiment Similarity Coefficients – 44-group Covariance Data



c(k) Color Key
1.0
0.8
0.6
0.4
0.2
0.0
-0.2

Benchmark Experiment Similarity Coefficients – 56-group Covariance Data

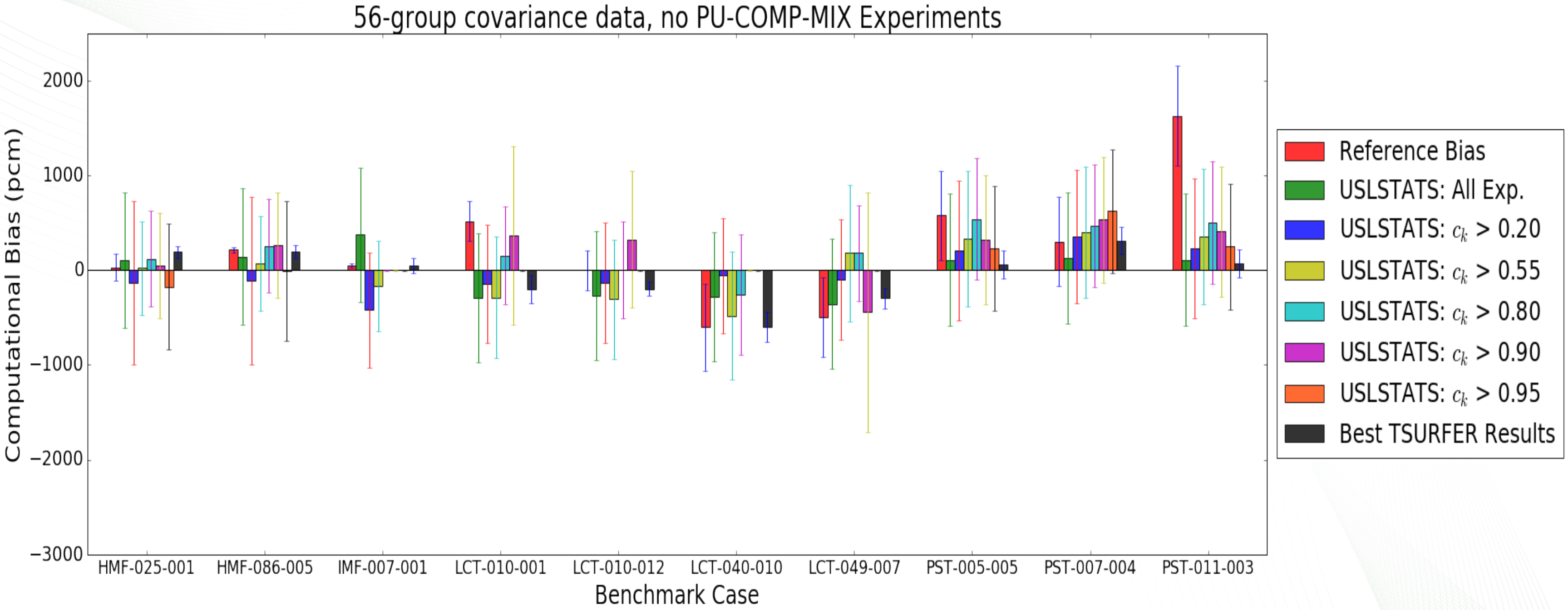
BFS Cases →
HEU-MET-FAST →
HEU-SOL-THERM →
IEU-MET-FAST →
LEU-COMP-THERM →
LEU-MET-THERM →
LEU-SOL-THERM →
MIX-COMP-THERM →
MIX-COMP-FAST →



← MIX-MISC-THERM
← MIX-SOL-THERM
← PU-COMP-MIXED
← PU-MET-FAST
← PU-SOL-THERM

c(k) Color Key
1.0
0.8
0.6
0.4
0.2
0.0
-0.2

Bias Results: Known Bias Cases USLSTATS c_k Trending Analysis



Max c_k	0.9986	0.9962	0.9642	0.9674	0.9694	0.9527	0.9766	0.9998	0.9999	0.9982
Cases with $c_k > 0.9$	14	14	1	10	16	2	4	55	56	55

USLSTATS Results – Takeaways:

< 2 σ Error
2-3 σ Error
> 3 σ Error

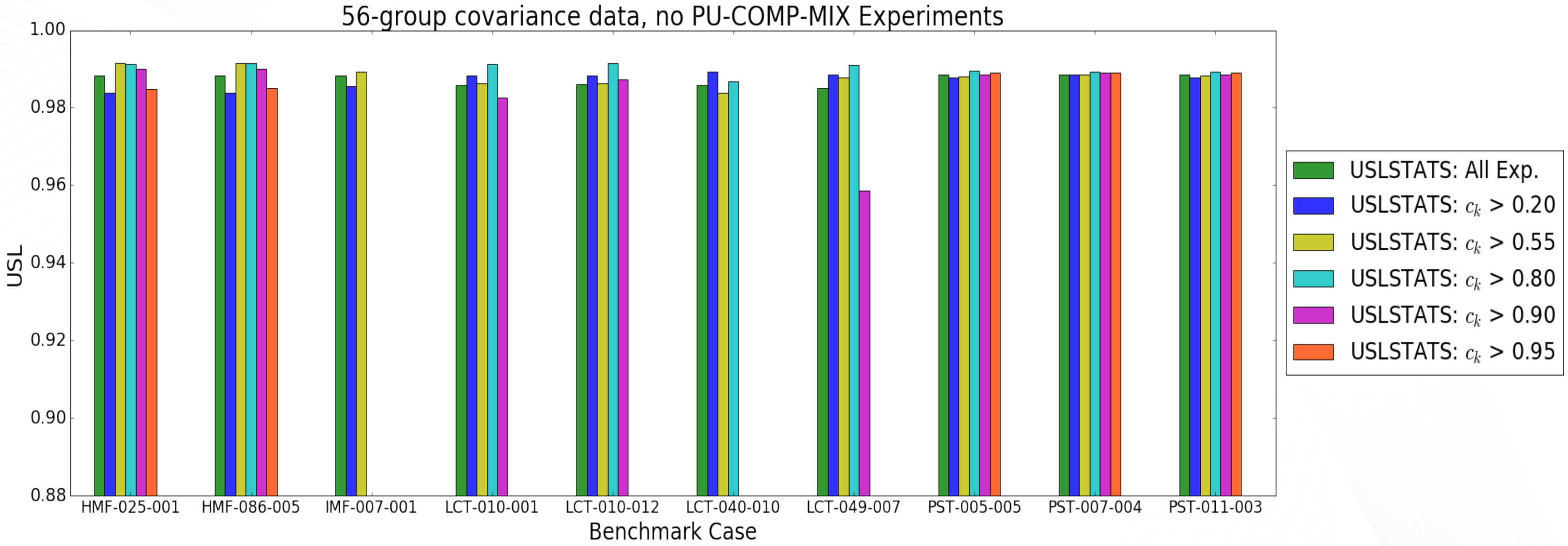
	Difference (Units of σ)	All Exp.	c_k > 0.2	c_k > 0.65 c_k > 0.55	c_k > 0.8	c_k > 0.9	c_k > 0.95	Best TSURFER Results
44-group Covariance Data	**Average**	**3.43**	**11.72**	**3.03**	**1.26***	**1.10***	**4.54***	**1.56**
	Max	**8.36**	**35.98**	**8.24**	**2.38***	**2.44***	**9.11***	**3.38**
56-group Covariance Data	**Average**	**5.23**	**7.93**	**3.86**	**1.20***	**1.17***	**3.46***	
	Max	**15.55**	**22.27**	**10.11**	**2.12***	**2.31***	**7.06***	

* Too few cases existed to compute bias estimates for at least one application.

USLSTATS Results – Takeaways:

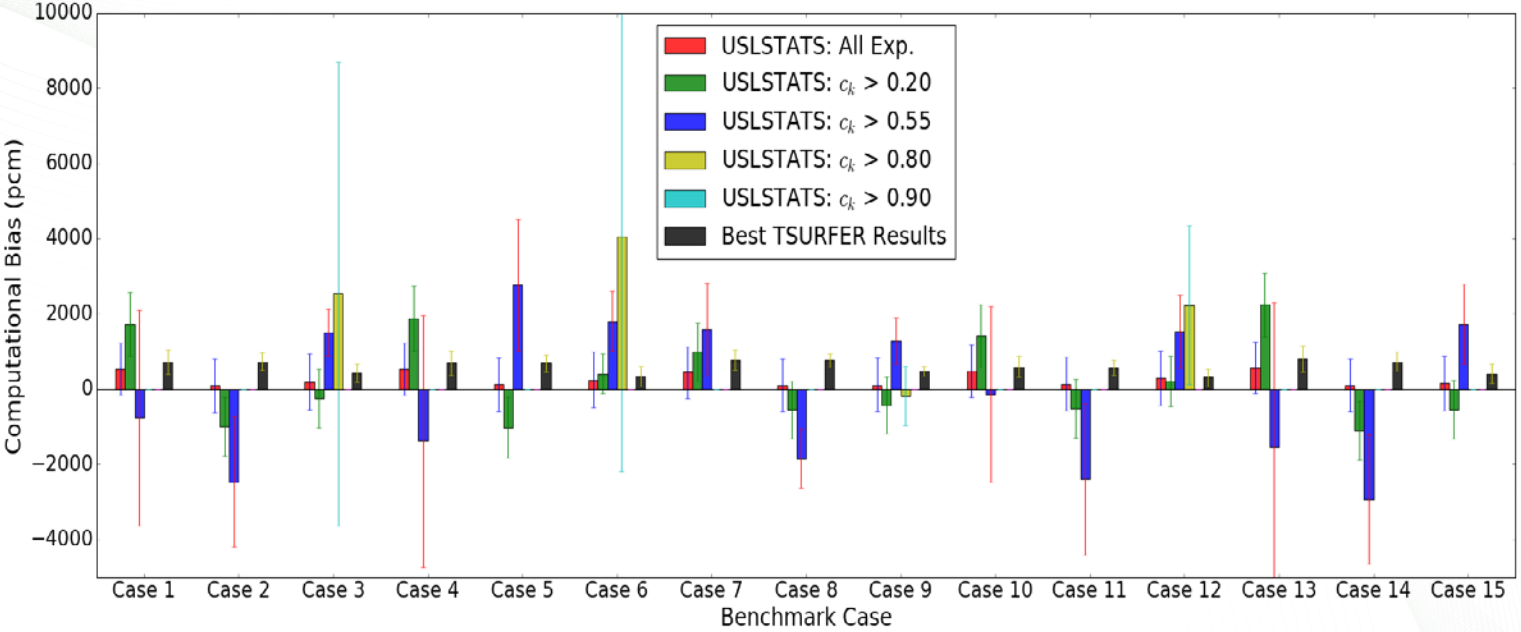
- USLSTATS predicted accurate computational biases when enough high similarity benchmark experiments were present.
- The “best” USLSTATS bias predictions were more accurate than the “best” TSURFER bias predictions.
- This comparison could be strengthened if the benchmark experiment results had smaller error bars.

USL Results: Known Bias Cases USLSTATS c_k Trending Analysis



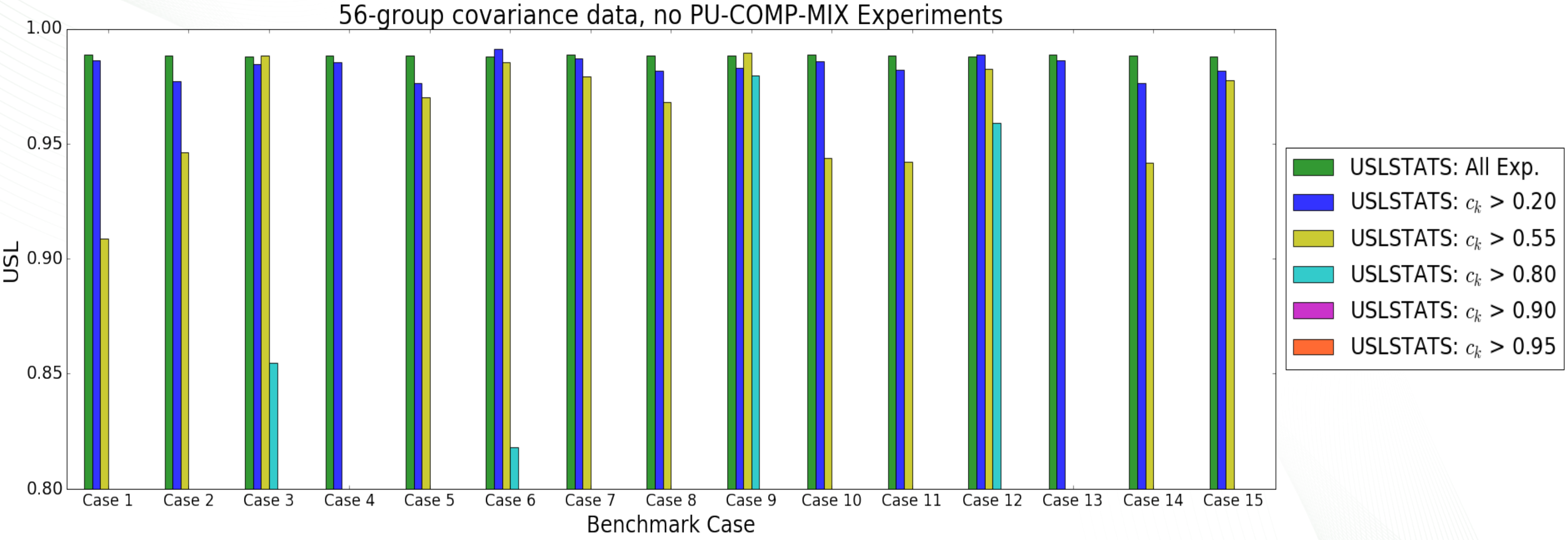
Max c_k	0.9986	0.9962	0.9642	0.9674	0.9694	0.9527	0.9766	0.9998	0.9999	0.9982
Cases with $c_k > 0.9$	14	14	1	10	16	2	4	55	56	55

Bias Results: Unknown Bias Cases USLSTATS c_k Trending Analysis



Max c_k	0.6935	0.6627	0.8370	0.6760	0.6822	0.8347	0.7753	0.7194	0.9293	0.7290	0.6683	0.8649	0.6566	0.6716	0.7950
Cases w/ $c_k > 0.9$	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

USL Results: Unknown Bias Cases USLSTATS c_k Trending Analysis



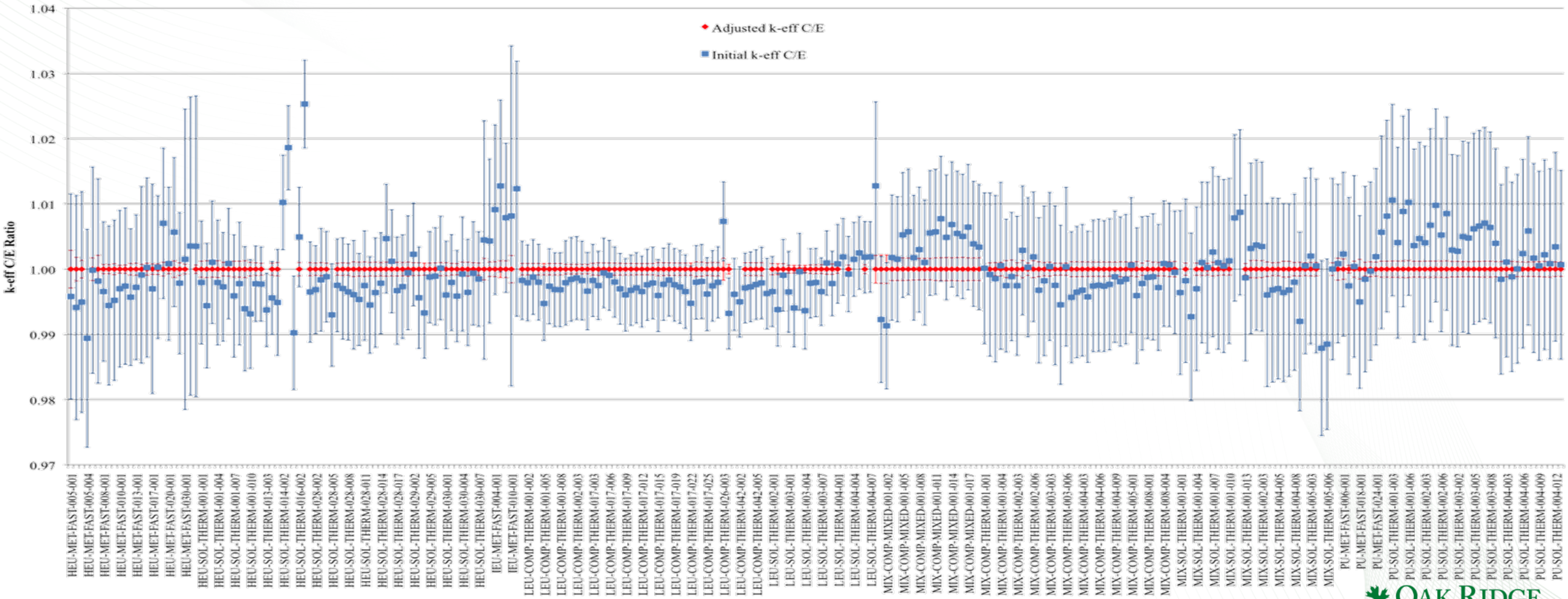
Max c_k	0.6935	0.6627	0.8370	0.6760	0.6822	0.8347	0.7753	0.7194	0.9293	0.7290	0.6683	0.8649	0.6566	0.6716	0.7950
Cases w/ $c_k > 0.9$	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

TSURFER Tools for Data Adjustment and Experimental Data Assimilation

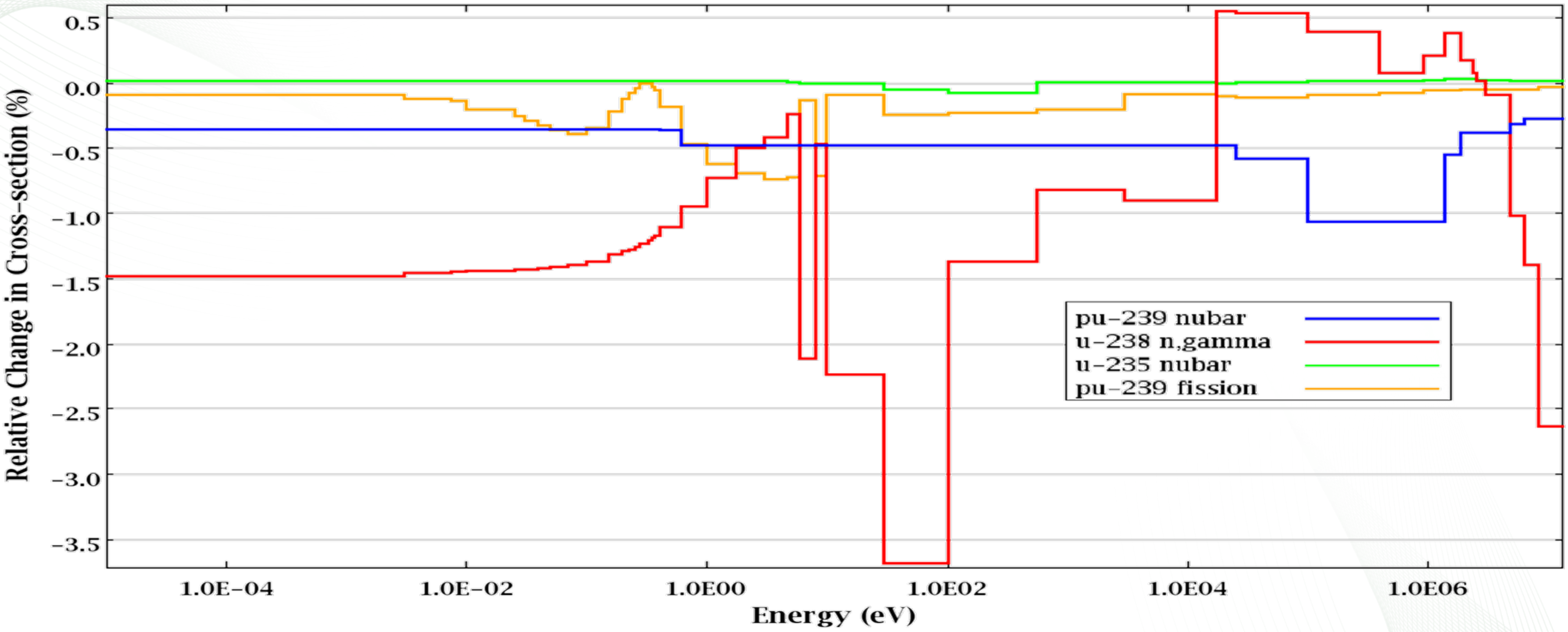
- *TSURFER:* *T*ool for *S*/*U* analysis of *R*esponse *F*unctionals using *E*xperimental *R*esults
  - Biases are observed as differences between benchmark and computed quantities (k_{eff}, reaction rates, etc.)
  - TSURFER uses sensitivity information to **consistently** adjust nuclear data and reconcile biases between integral experiment results and computational predictions.
  - Where the cross sections and covariance data are modified, the modifications can be used to project biases from the benchmarks to targeted application systems.

Data Adjustment Techniques:

Experimental benchmark data (E) is used to improve the accuracy of the initial computed responses (C).



TSURFER Cross Section Adjustments



OAK RIDGE National Laboratory

A Note on Bias

- The computational bias measures the predictive capabilities of a modeling and simulation tool.
- For USLSTATS:

$$Relative\ Bias = C/E_{Extrapolated} - 1$$

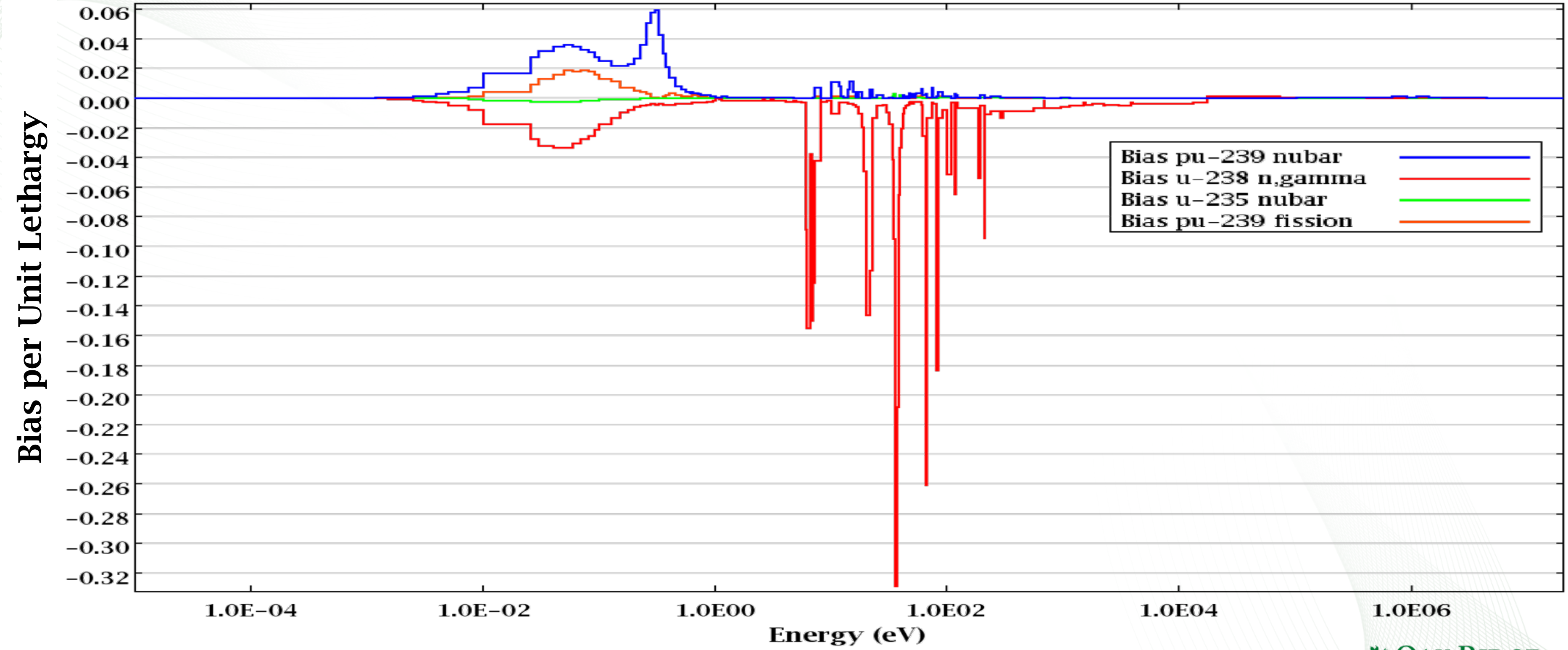
- For TSURFER:

$$Relative\ Bias = \frac{(Calculated\ Response\ - Adjusted\ Response)}{Calculated\ Response}$$

OAK RIDGE National Laboratory

Bias Estimation

$$\textbf{Bias} = k_{\text{Simulation}} - k_{\text{Experiment}}$$



Adjusted Cross Sections Reduce Data-Induced Biases

- Original Application Uncertainty is:
  0.520% Δk/k

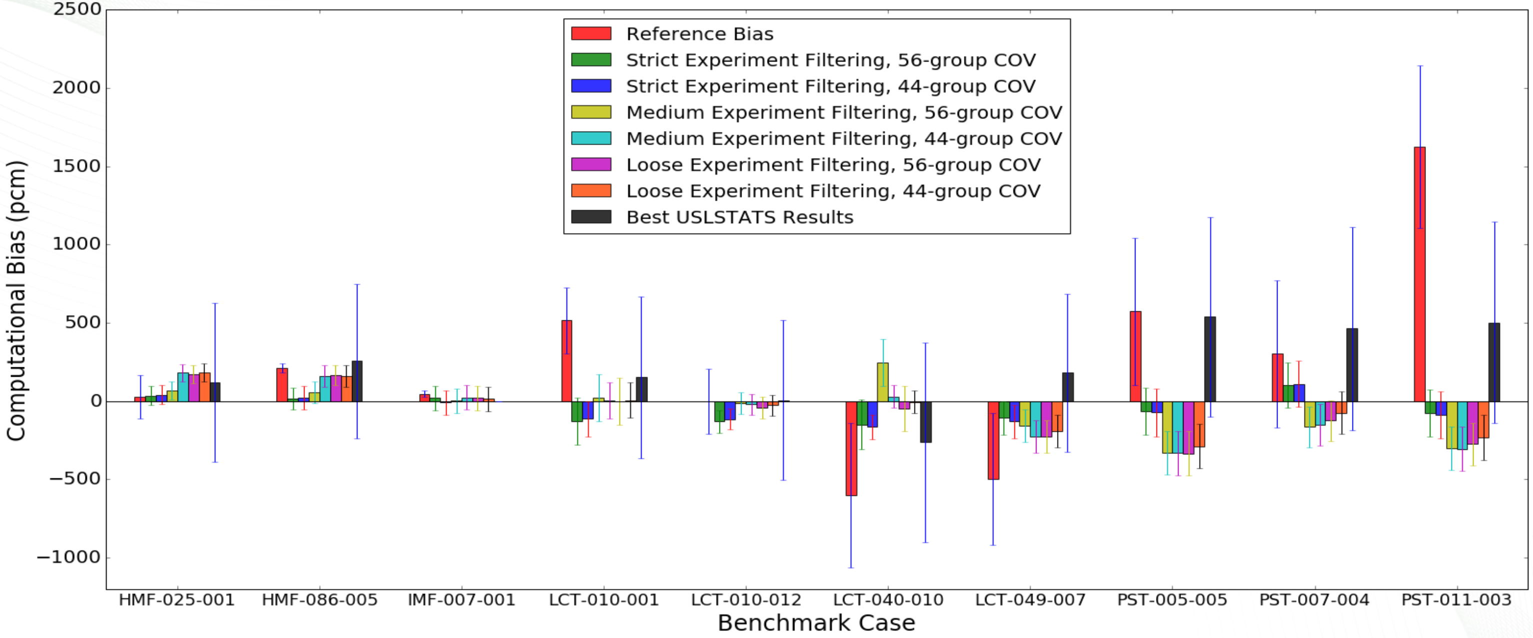
- Adjusted Application Uncertainty is:
  0.119% Δk/k

- Interpretation: ~80% of uncertainty is quantified through validation with experiments.

- Remaining uncertainty highlights gaps in available validation data.

NUCLIDE	REACTION	CONTRIBUTION TO BIAS % dk/k
u-238	n,gamma	-2.1084E-01
pu-239	nubar	1.2761E-01
pu-239	fission	3.9872E-02
o-16	elastic	3.2243E-02
pu-239	n,gamma	-2.5810E-02
pu-239	chi	1.0248E-02
u-235	chi	2.9940E-04
fe-56	n,gamma	1.7158E-02
u-235	fission	-1.2351E-02
pu-240	n,gamma	-1.3162E-02
u-238	elastic	2.7715E-03
u-235	n,gamma	1.0599E-03
h-1	elastic	2.7348E-03
u-238	n,n'	-6.8963E-03
u-235	nubar	-4.1298E-03
fe-56	elastic	-6.0079E-03
h-1	n,gamma	4.1893E-03
u-238	nubar	3.1408E-03

Bias Results: Known Bias Cases TSURFER Analysis



Max c_k	0.9986	0.9962	0.9642	0.9674	0.9694	0.9527	0.9766	0.9998	0.9999	0.9982
Cases with $c_k > 0.9$	14	14	1	10	16	2	4	55	56	55

Oak Ridge National Laboratory

TSURFER Results – Takeaways:

< 2 σ Error
2-3 σ Error
> 3 σ Error

	Difference (Units of σ)	Strict Filtering	Medium Filtering	Loose Filtering	Best USLSTATS Results
44-group Covariance Data	**Average**	**3.04**	**1.61**	**1.59**	**1.20***
	Max	**-8.13**	**-3.03**	**-3.04**	**2.12***
56-group Covariance Data	**Average**	**3.39**	**1.56**	**1.53**	
	Max	**-9.59**	**-3.38**	**-3.35**	

* Too few cases existed to compute bias estimates for at least one application.

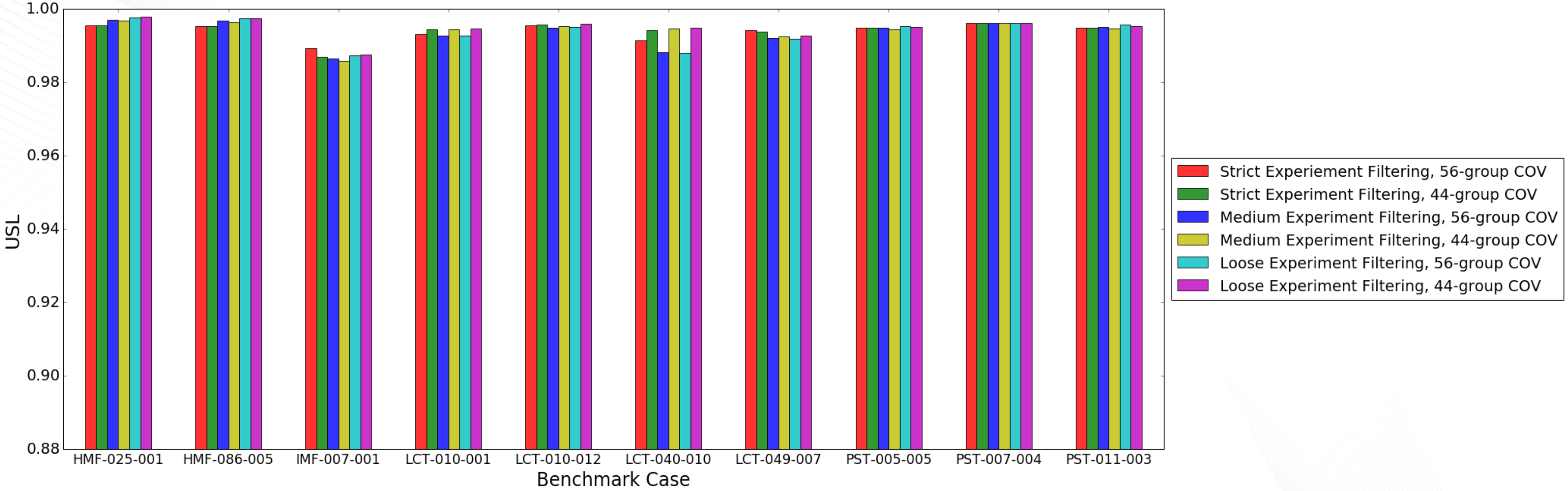
OAK RIDGE National Laboratory

TSURFER Results – Takeaways:

- TSURFER bias predictions were slightly less accurate than the best USLSTATS bias predictions.

- The TSURFER bias predictions were significantly more consistent than the USLSTATS predictions (more discussion about this later).

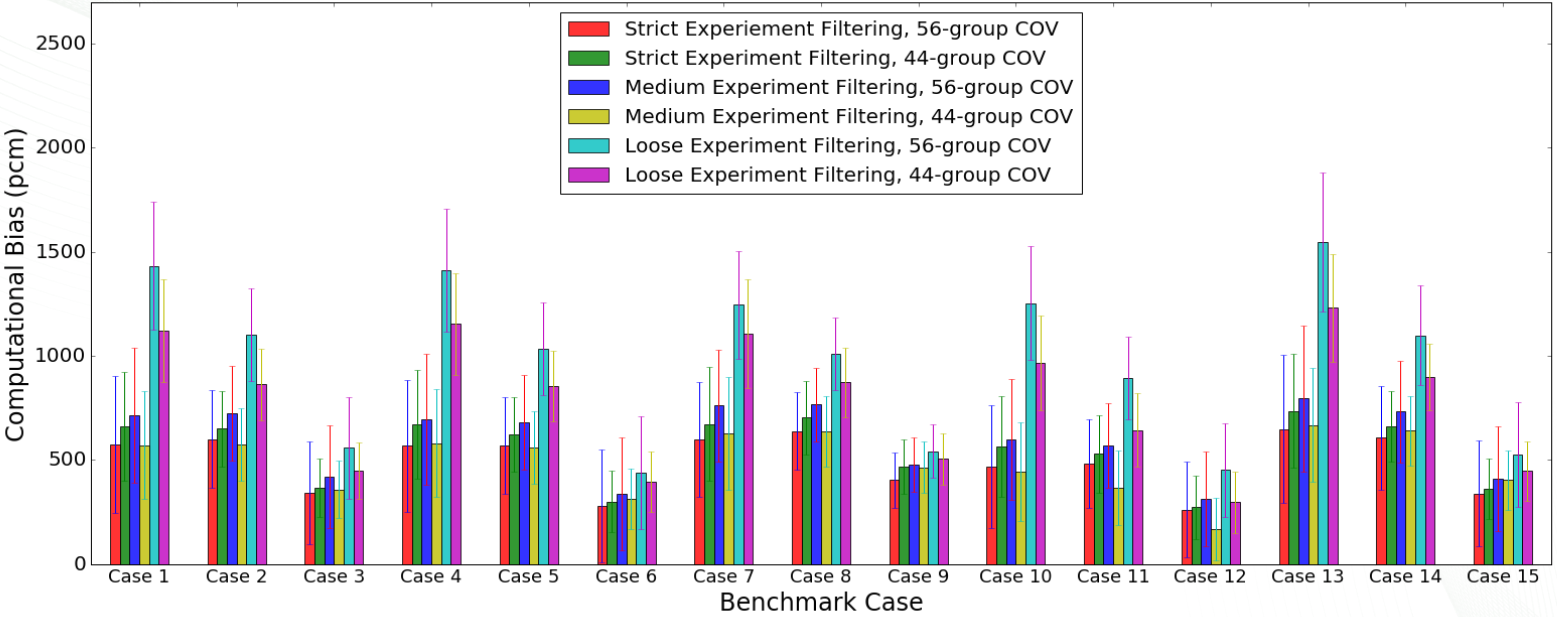
USL Results: Known Bias Cases TSURFER Analysis



Max c_k	0.9986	0.9962	0.9642	0.9674	0.9694	0.9527	0.9766	0.9998	0.9999	0.9982
Cases with $c_k > 0.9$	14	14	1	10	16	2	4	55	56	55

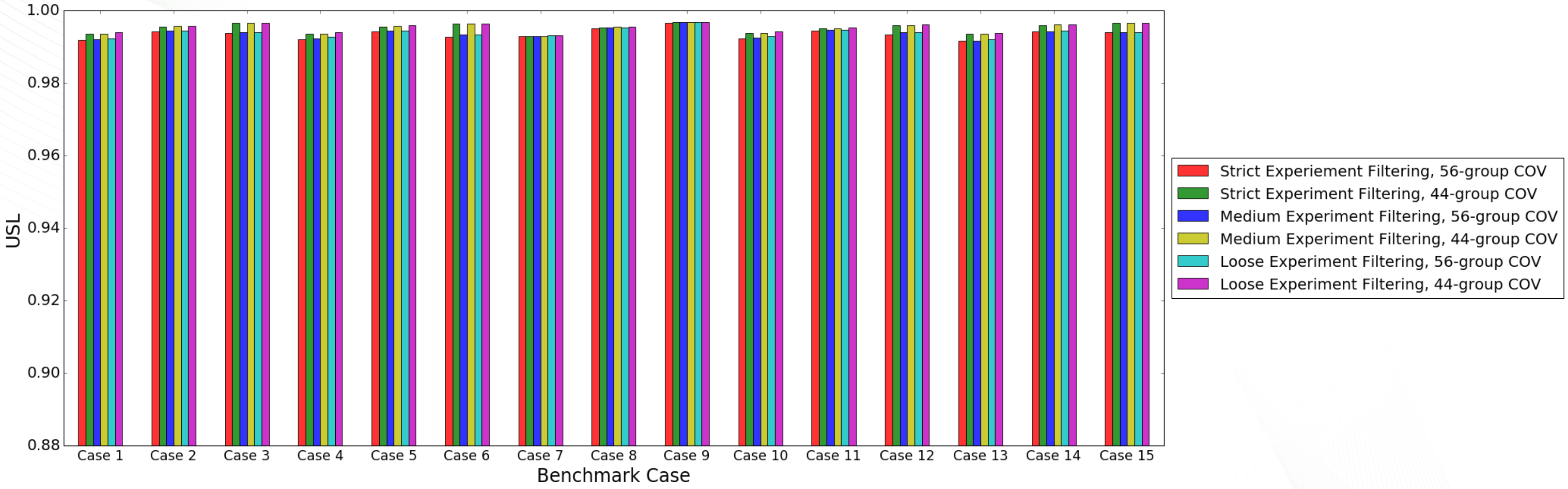
OAK RIDGE National Laboratory

USL Results: Unknown Bias Cases TSURFER Data Assimilation Analysis



Max c_k	0.6935	0.6627	0.8370	0.6760	0.6822	0.8347	0.7753	0.7194	0.9293	0.7290	0.6683	0.8649	0.6566	0.6716	0.7950
Cases w/ $c_k > 0.9$	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

USL Results: Unknown Bias Cases TSURFER Data Assimilation Analysis



Max c_k	0.6935	0.6627	0.8370	0.6760	0.6822	0.8347	0.7753	0.7194	0.9293	0.7290	0.6683	0.8649	0.6566	0.6716	0.7950
Cases w/ c_k > 0.9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

TSURFER Results – Takeaways:

- TSURFER bias predictions were slightly less accurate than the best USLSTATS bias predictions.
- The TSURFER bias predictions were significantly more consistent than the USLSTATS predictions (more discussion about this later).
- TSURFER USL estimates are much closer to 1.00 than USLSTATS estimates.

TSURFER: ~0.99 **USLSTATS:** 0.96 – 0.98

Whisper Methodology

- Non-parametric, extreme-value theory based method for determining USL's and margins of subcriticality (MOS).
  - Developed by Kiedrowski at LANL in 2014
- Whisper USL's are designed to be conservative.
  - Adding more benchmark cases can only increase the MOS.
- Results presented here were obtained using an independently-developed implementation of the Whisper methodology.

Whisper Methodology

- Whisper weights the importance of experiments based on their similarity to the target application.

$$weight_i = \frac{\left(c_k^i - c_k^{threshold}\right)}{\left(\max(c_k) - c_k^{threshold}\right)}$$

- Whisper includes additional benchmark experiments until a cumulative weight of 25 is obtained.
  - Treatments exist for applying the Whisper method to cases with few similar benchmarks.

Whisper Methodology

- Whisper uses data assimilation methods to identify inconsistent benchmark experiments and omit them from the USL calculation.
- Whisper uses the adjusted response uncertainty to provide additional subcritical margin.
  - Performing a convergence study on the adjusted response uncertainty is helpful for Whisper analyses.

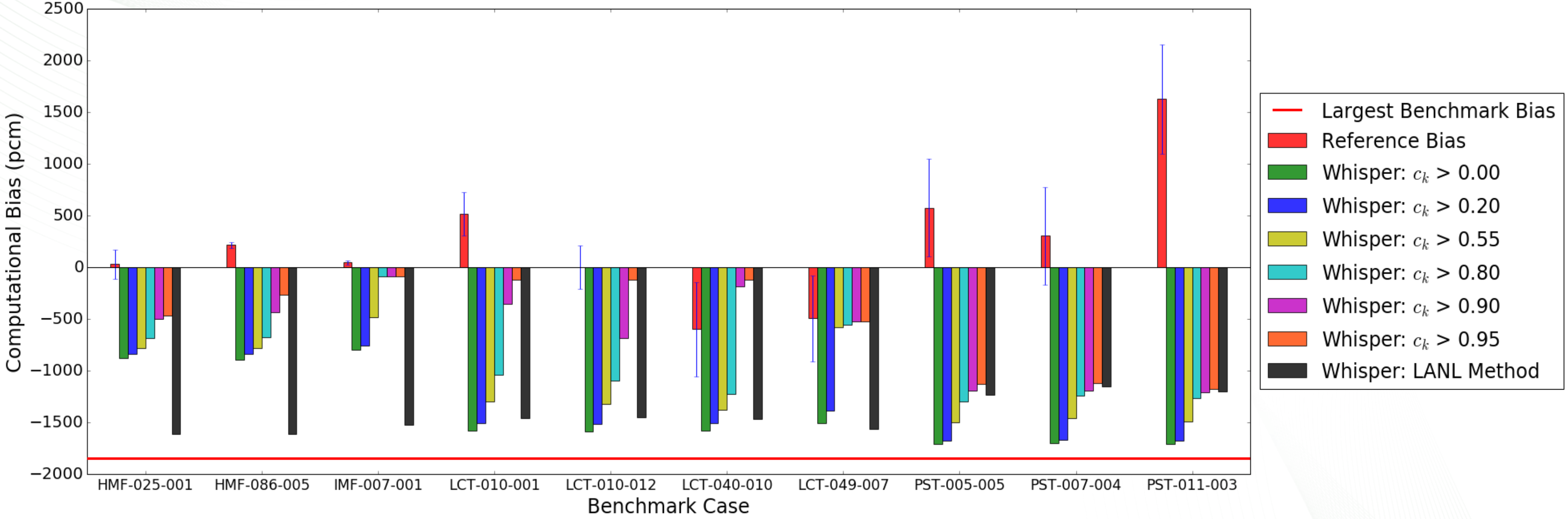
$$MOS = MOS_{software} + MOS_{data} + MOS_{application}$$

- A detailed discussion of the Whisper methodology is available in:

B.C. Kiedrowski, et. al., "Whisper: Sensitivity/Uncertainty-Based Computational Methods and Software for Determining Baseline Upper Subcritical Limits," *Nucl. Sci. Eng.* (2015).

OAK RIDGE National Laboratory

Bias Results: Known Bias Cases Whisper Implementation Analysis



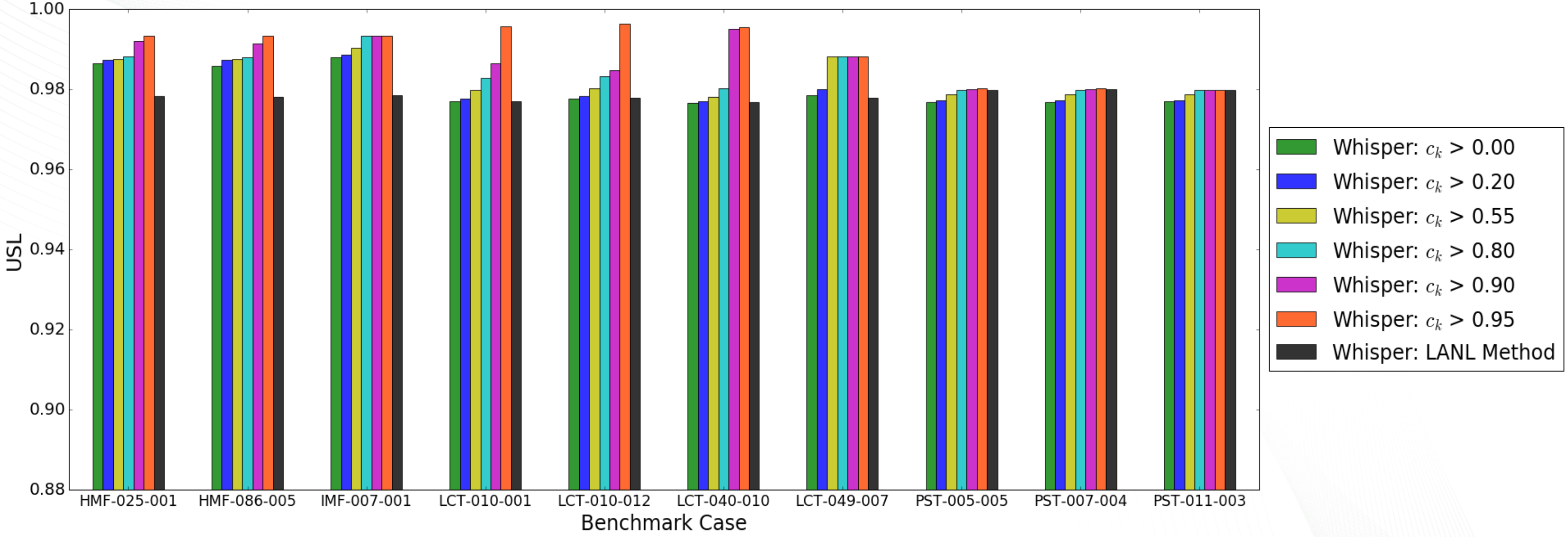
Max c_k	0.9986	0.9962	0.9642	0.9674	0.9694	0.9527	0.9766	0.9998	0.9999	0.9982
Cases with $c_k > 0.9$	14	14	1	10	16	2	4	55	56	55

Whisper Results – Takeaways:

- Whisper bias estimates should be interpreted differently than USLSTATS or TSURFER bias estimates.
- When few highly similar experiments exist, the Whisper bias approaches the bias of the most conservative experiment available.
- As the c_k threshold decreases, the Whisper bias approaches the bias of the most conservative experiment available.

OAK RIDGE National Laboratory

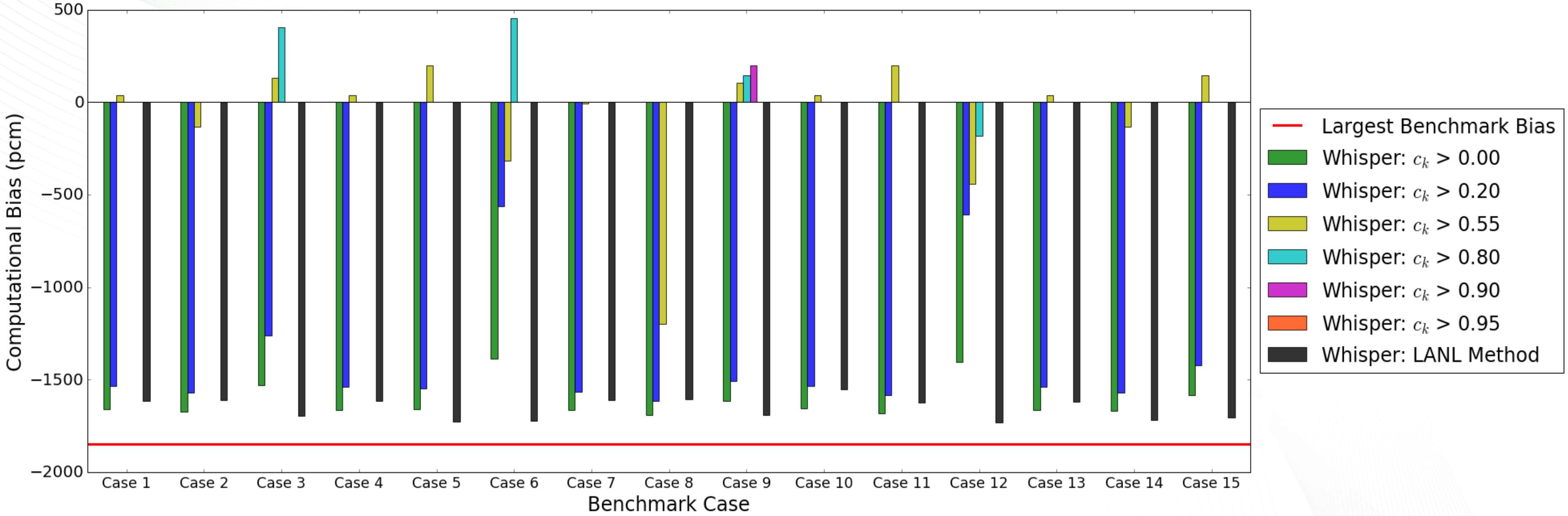
USL Results: Known Bias Cases Whisper Implementation Analysis



Max c_k	0.9986	0.9962	0.9642	0.9674	0.9694	0.9527	0.9766	0.9998	0.9999	0.9982
Cases with $c_k > 0.9$	14	14	1	10	16	2	4	55	56	55

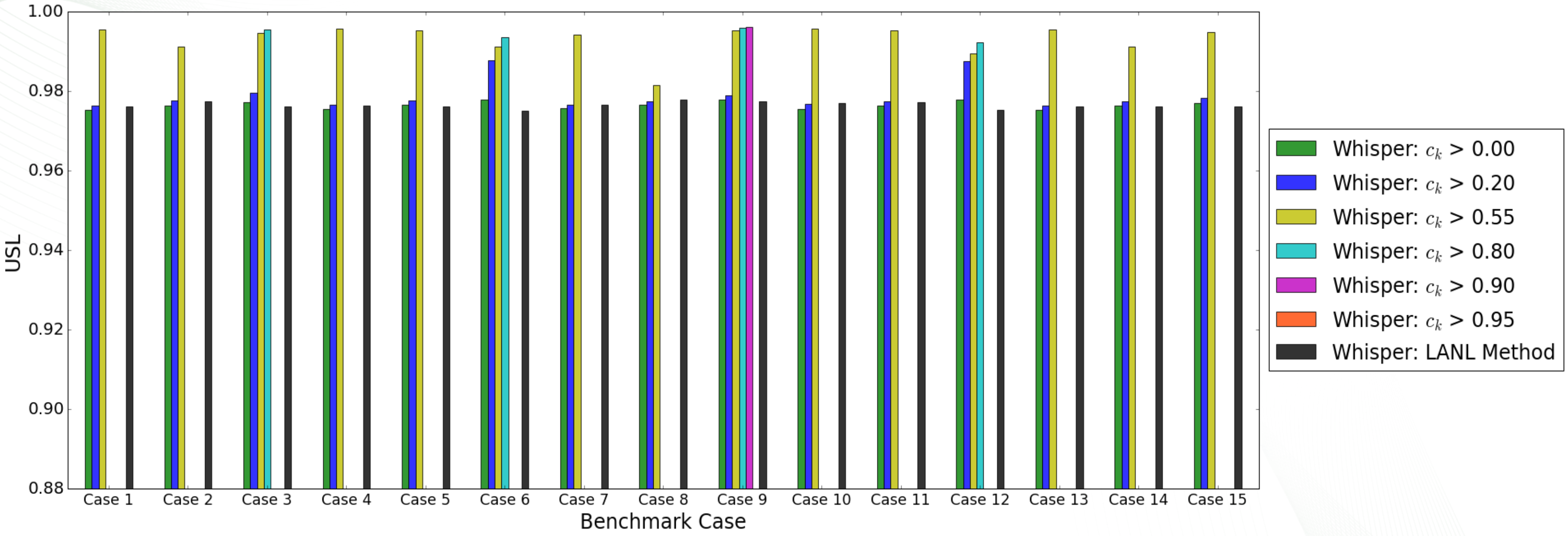
OAK RIDGE National Laboratory

Bias Results: Unknown Bias Cases Whisper Implementation Analysis



Max c_k	0.6935	0.6627	0.8370	0.6760	0.6822	0.8347	0.7753	0.7194	0.9293	0.7290	0.6683	0.8649	0.6566	0.6716	0.7950
Cases w/ c_k > 0.9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

USL Results: Unknown Bias Cases
Whisper Implementation Analysis



Max c_k	0.6935	0.6627	0.8370	0.6760	0.6822	0.8347	0.7753	0.7194	0.9293	0.7290	0.6683	0.8649	0.6566	0.6716	0.7950
Cases w/ $c_k > 0.9$	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

Standard Deviation of Bias Estimates

Covariance Data	Application Cases	USLSTATS	TSURFER (all cases)	TSURFER (cases with filtering)	Whisper (all cases)	Whisper (cases with at least w_i)
44-group	Phase I	361	121	122	441	168
	Phase V	767	162	81	802	50
56-group	Phase I	231	169	166	455	172
	Phase V	1,338	240	101	819	110
Overall	Phase I	295	142	136	434	173
	Phase V	1,020	199	87	1,167	88

* Standard deviations expressed in units of pcm.

Standard Deviation of Bias Estimates

Covariance Data	Application Cases	USLSTATS	TSURFER (all cases)	TSURFER (cases with filtering)	Whisper (all cases)	Whisper (cases with at least w_i)
44-group	Phase I	361	121	122	441	168
	Phase V	767	162	81	802	50
56-group	Phase I	231	169	166	455	172
	Phase V	1,338	240	101	819	110
Overall	Phase I	295	142	136	434	173
	Phase V	1,020	199	87	1,167	88

* Standard deviations expressed in units of pcm.

OAK RIDGE National Laboratory

Standard Deviation of USL Estimates

Covariance Data	Application Cases	USLSTATS	TSURFER (all cases)	TSURFER (cases with filtering)	Whisper (all cases)	Whisper (cases with at least w_i)
44-group	Phase I	1,423	115	119	481	115
	Phase V	7,554	61	65	773	70
56-group	Phase I	522	140	137	504	117
	Phase V	4,628	77	83	806	79
Overall	Phase I	1,088	140	145	485	120
	Phase V	6,431	121	114	771	79

* Standard deviations expressed in units of pcm.

Standard Deviation of USL Estimates

Covariance Data	Application Cases	USLSTATS	TSURFER (all cases)	TSURFER (cases with filtering)	Whisper (all cases)	Whisper (cases with at least w_i)
44-group	Phase I	1,423	115	119	481	115
	Phase V	7,554	61	65	773	70
56-group	Phase I	522	140	137	504	117
	Phase V	4,628	77	83	806	79
Overall	Phase I	1,088	140	145	485	120
	Phase V	6,431	121	114	771	79

* Standard deviations expressed in units of pcm.

TSURFER Results – Takeaways:

- USLSTATS bias and USL estimates can vary significantly based on the user's input parameters and choice of benchmark experiments.

- The TSURFER and Whisper bias estimates exhibited the greatest degree of consistency.

Conclusions

- USLSTATS and TSURFER produced accurate bias estimates for cases with known biases.
  - USLSTATS trending analyses were found to be most effective when trending on c_k.
  - Setting a very high c_k threshold CAN decrease the accuracy of USLSTATS.
  - High uncertainty in the known bias reference cases makes it difficult to evaluate the accuracy of the bias prediction methods.
- TSURFER and Whisper produced consistent bias and USL estimates for cases with unknown biases, but USLSTATS did not.
- The TSURFER USL estimates are generally less conservative than those from USLSTATS or Whisper.
  - TSURFER was designed for accurate prediction of bias, not conservative USL estimates.

Conclusions

- USLSTATS and TSURFER produced accurate bias estimates for cases with known biases.
  - USLSTATS trending analyses were found to be most effective when trending on c_k.
  - Setting a very high c_k threshold CAN decrease the accuracy of USLSTATS.
  - High uncertainty in the known bias reference cases makes it difficult to evaluate the accuracy of the bias prediction methods.
- TSURFER and Whisper produced consistent bias and USL estimates for cases with unknown biases, but USLSTATS did not.
- The TSURFER USL estimates are generally less conservative than those from USLSTATS or Whisper.
  - TSURFER was designed for accurate prediction of bias, not conservative USL estimates.

Conclusions

- USLSTATS and TSURFER produced accurate bias estimates for cases with known biases.
  - USLSTATS trending analyses were found to be most effective when trending on c_k.
  - Setting a very high c_k threshold CAN decrease the accuracy of USLSTATS.
  - High uncertainty in the known bias reference cases makes it difficult to evaluate the accuracy of the bias prediction methods.
- TSURFER and Whisper produced consistent bias and USL estimates for cases with unknown biases, but USLSTATS did not.
- The TSURFER USL estimates are generally less conservative than those from USLSTATS or Whisper.
  - TSURFER was designed for accurate prediction of bias, not conservative USL estimates.