Reducing uncertainties in chosen neutron cross-sections through the use of Approximate Bayesian Computation

Supervisor: prof. Tomasz Kozłowski
Co-supervisor: dr. Piotr Kopka

Michał Jędrzejczyk

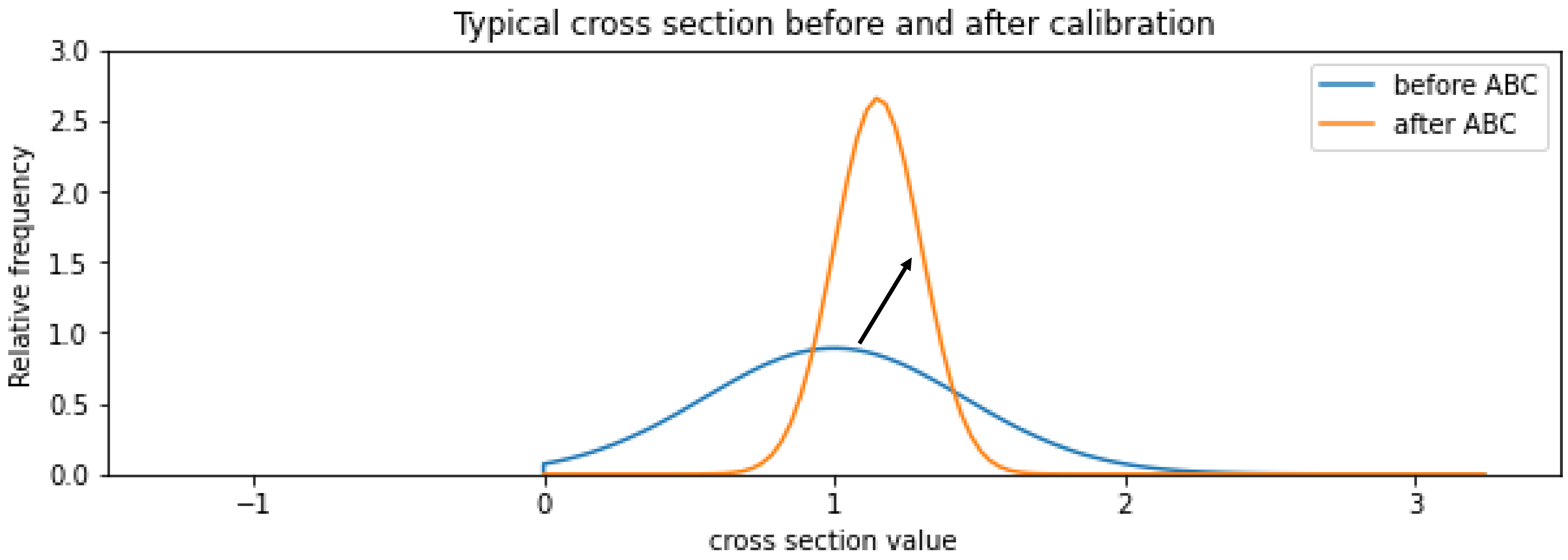
NARODOWE CENTRUM BADAŃ JĄDROWYCH ŚWIERK

Presentation plan

- Goal of my research
- Nuclear data deficiencies
- Bayesian Statistics and Approximate Bayesian Computation (ABC)
- Example of use of ABC
- Using ABC to improve nuclear data
- State of my research

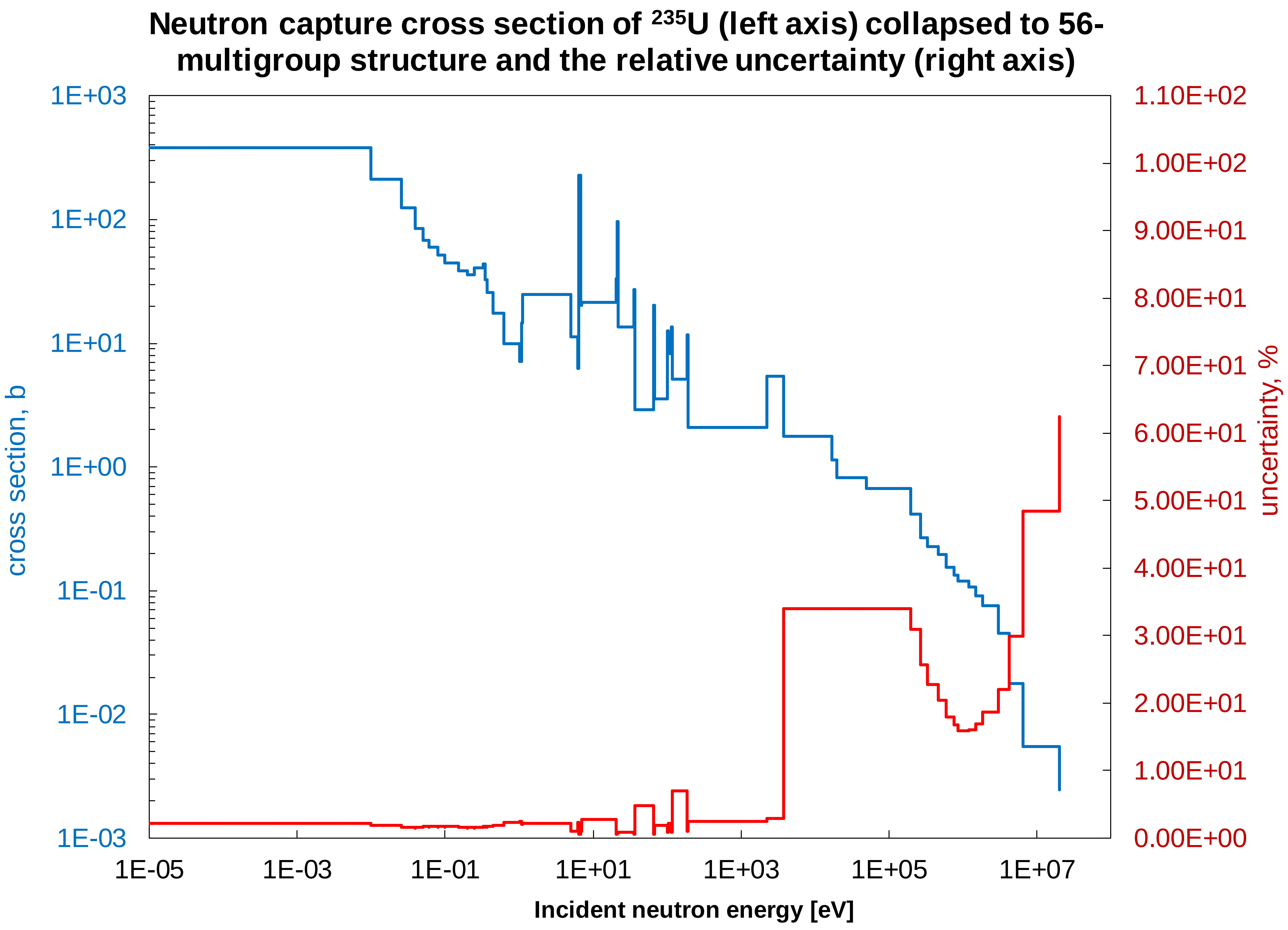
What do I want to achieve with my research?

Reduction of uncertainty in chosen neutron cross sections for ^{235}U in a library collapsed to 56-multigroup structure



This in turn will decrease uncertainty in computed multiplication factor (keff) value.

Inaccuracies in keff calculations



Cross section uncertainties are the main contributor to the keff uncertianty.

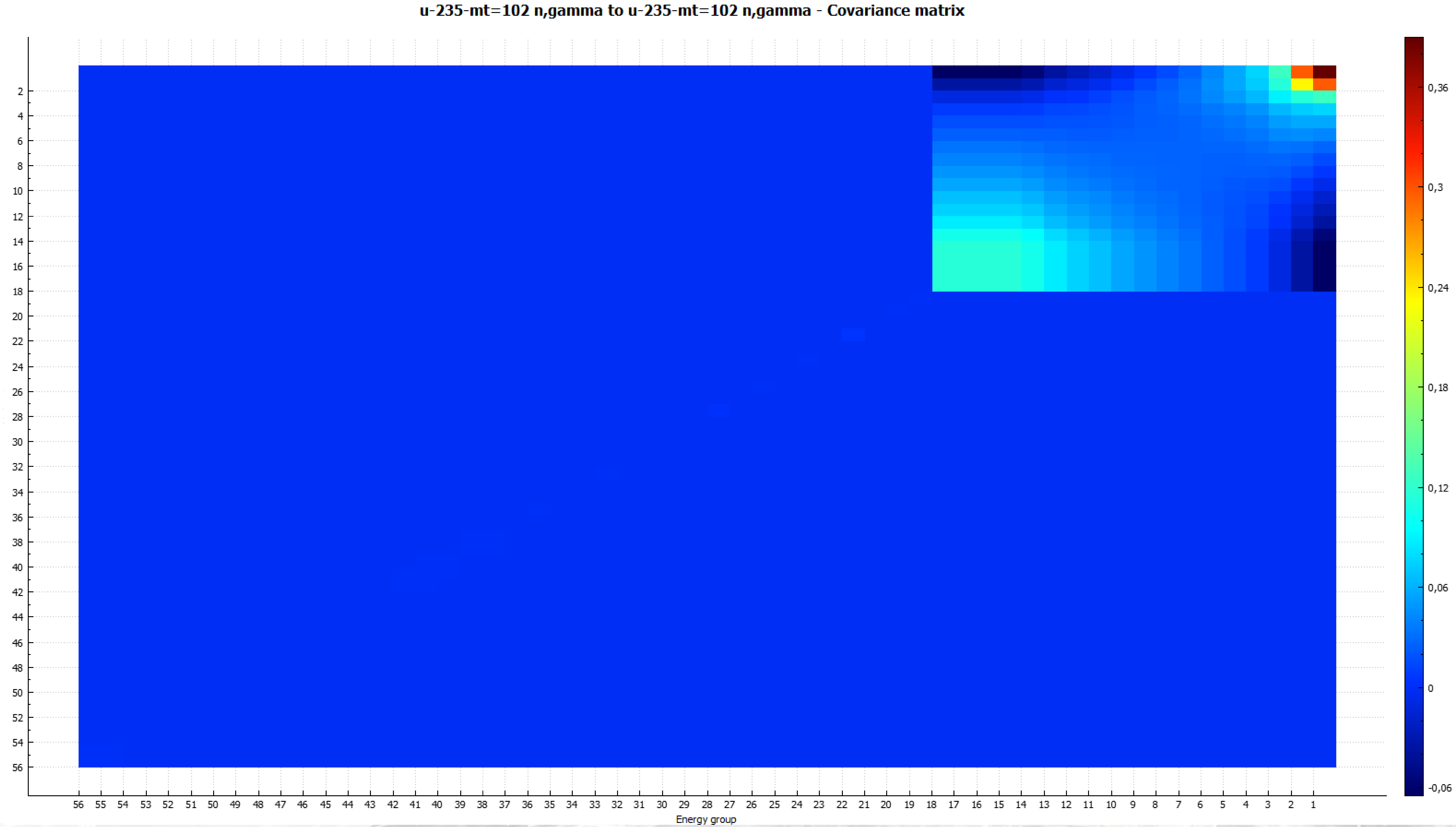
Keff uncertainty can be calculated with error propagation method. Most common way: calculate sensitivity to each reaction (for each energy range), and use the following equation:

$$E^2 = SMS^T$$

Where E = uncertainty of keff in 1 standard deviation
S = sensitivity vector
M = covariance matrix of xs uncertainties

Usually keff uncertainty ≈ 1-2%

Covariance matrix example



Is 1 % uncertainty a lot?

Prompt criticality for fast systems: keff=1.0026
safety margin: keff<1.0013

Result: suboptimal reactor geometry → deteriorated economics

The majority of uncertainty comes from the minority of neutron cross sections – for pure highly enriched fast uranium systems calculated with 56-multigroup library 75 % of uncertainty comes from just 23 cross sections

How to decrease uncertainty in model parameters (cross-sections in this case)?

It may be possible to decrease the uncertainty by the use of Bayesian Statistics.

$$\text{Bayes}'\text{theorem: } P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

$P(\theta)$ - prior (initial belief about probability or distribution of θ)

$P(\theta|D)$ - posterior (probability or distribution of θ given data D)

$P(D|\theta)$ - likelihood (joint probability of the data D as a function of θ)

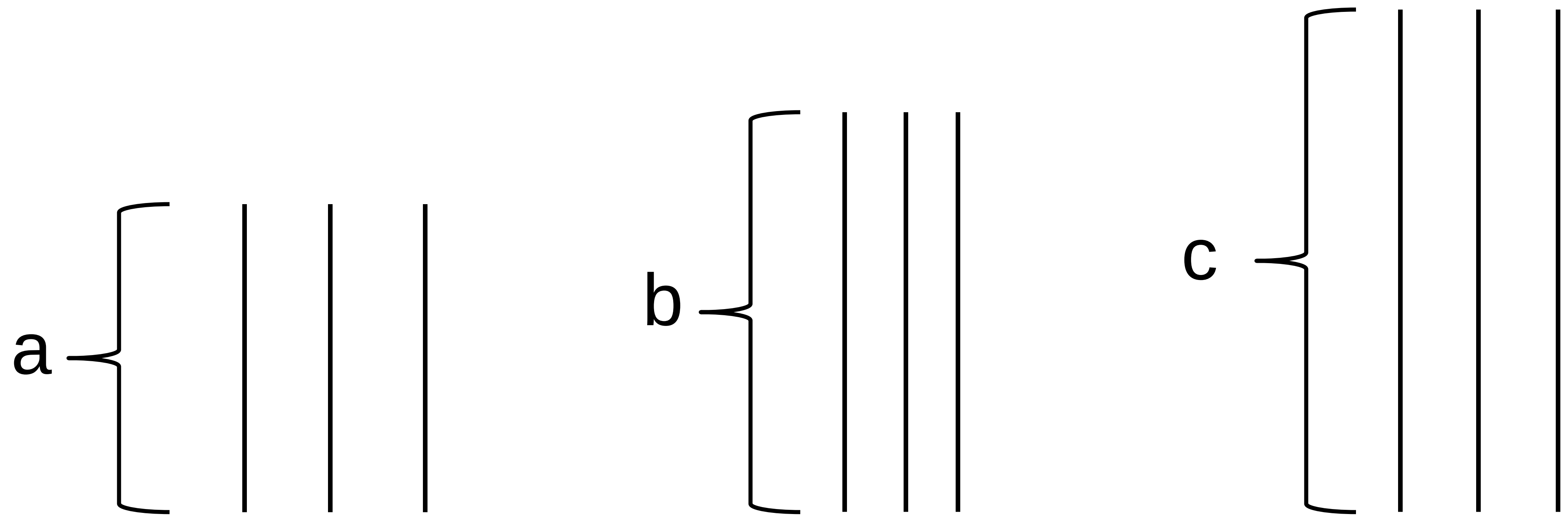
$P(D)$ - probability of data under any parameters

Finding distribution of model parameters

Given mathematical model $F = F(\theta)$, where θ are model parameters we want to improve the accuracy with which we know θ. Let's start with something simple:

a, b and c are rod lengths known to an experimenter with limited accuracy given by one standard deviation:
$a = 28 \pm 5\ cm$,
$b = 43 \pm 3\ cm$,
$c = 47 \pm 2\ cm$,
while their unknown true lengths are 30, 40 and 50 cm respectively.

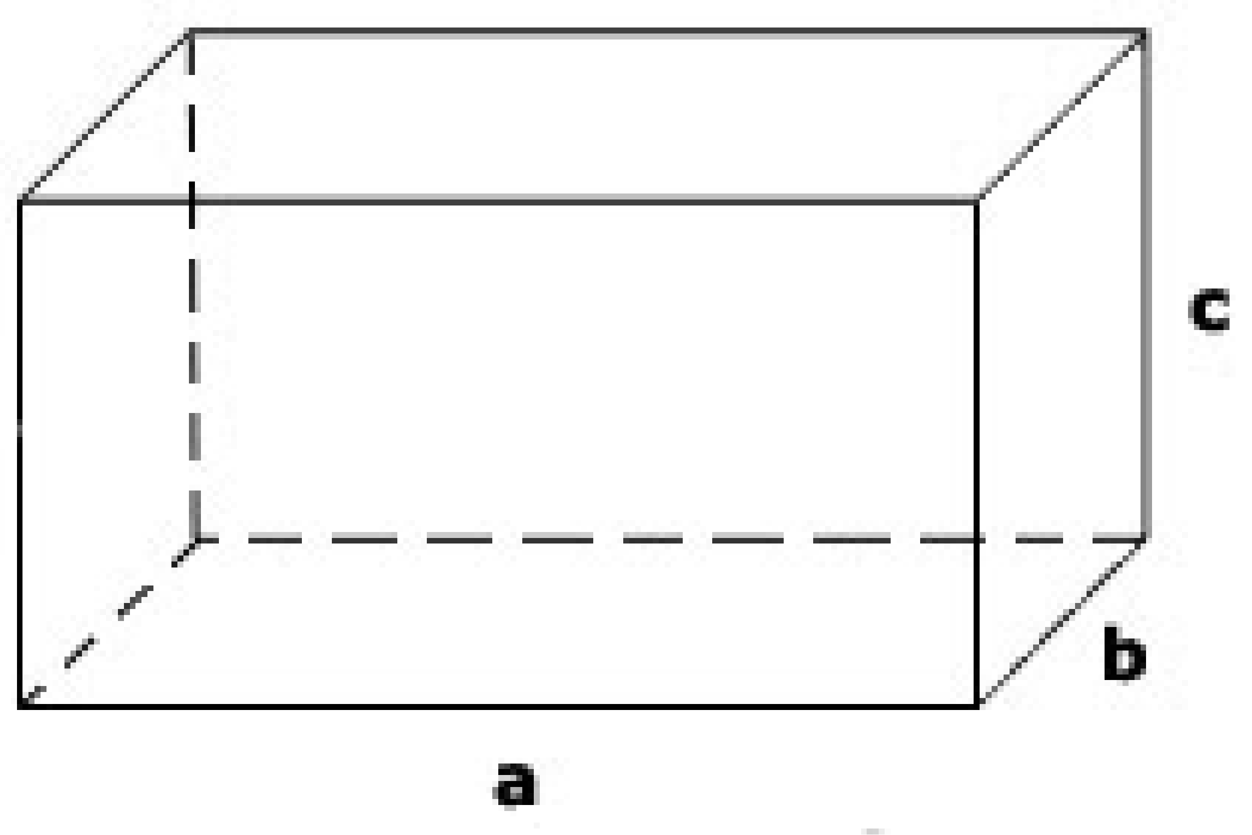
The experimenter wonders how he can improve his knowledge of rod lengths.



ABC example

The experimenter figures out, that he has a vary good weight measurement device. He constructs a cuboid out of the rods, adds walls to it, fills with water and measures weight to determine the volume. The formula of volume given rod lengths: $F = a * b * c$

Where F is a volume of a cuboid.



The measurement finds that the cuboid volume is 60030 $\pm$ 20 cm^3 (true value is 60000 cm^3).

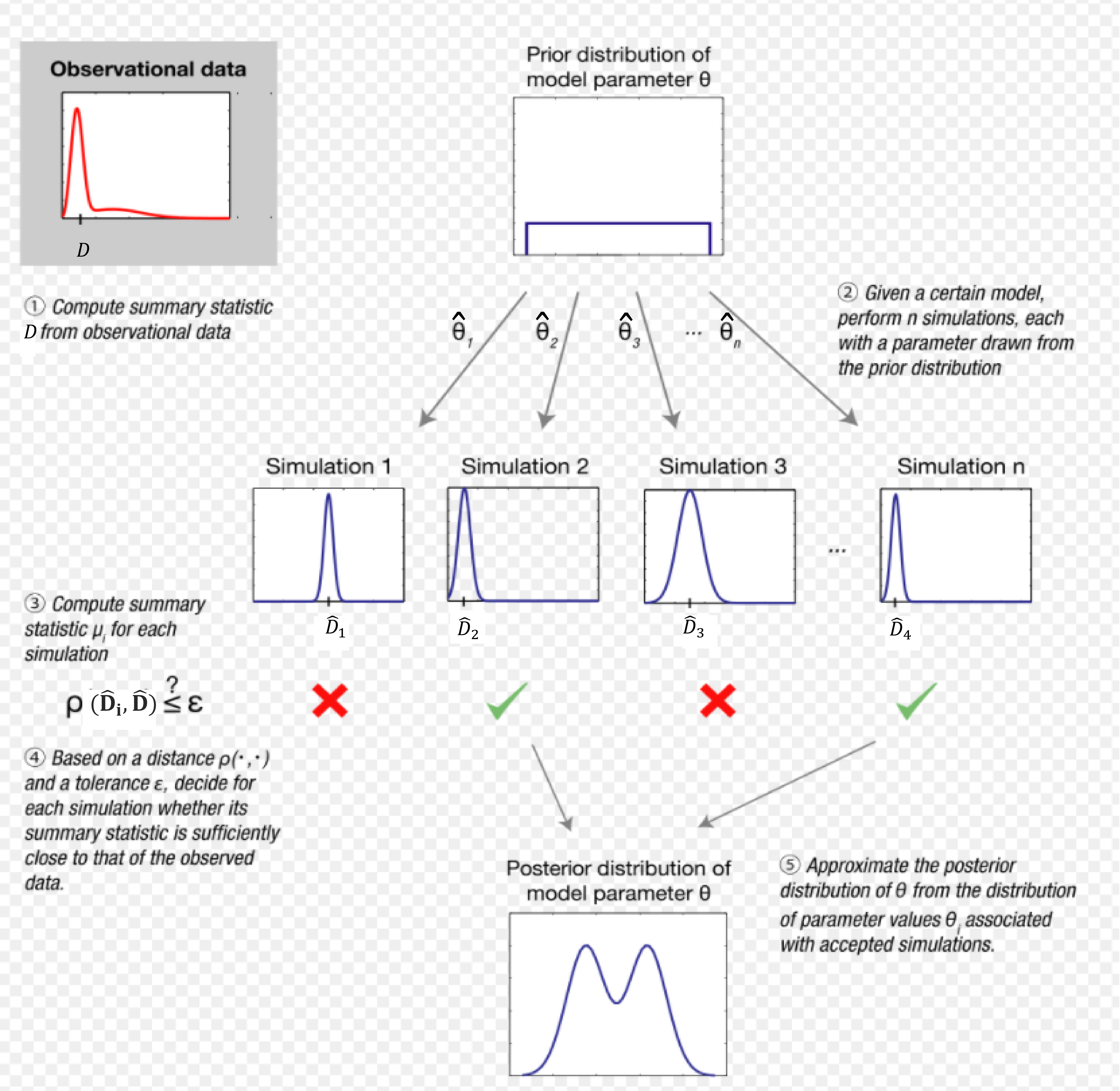
By an error propagation method, the experimenter estimates cuboid volume F calculated from $F = a * b * c$ is 56469 $\pm$ 11198 cm^3.

We can see intuitively that we should be able to extract values of a, b and c with better accuracy based on that measurement. But how?

How to find distribution of model parameters? - Approximate Bayesian Computation (ABC)

ABC estimates the posterior of a parameter θ by simulating the model to produce artificial data sets $\hat{D}$ using sample parameters $\hat{\theta}$ taken from the prior distribution. A comparison of how similar the artificial data set produced by the sample parameter from the prior to the observed real data set D is undertaken[1].

So for example from previous slide $\hat{\theta} = [\hat{a}, \hat{b}, \hat{c}]$ and we sample from distributions: $a = 28 \pm 5\ cm$, $b = 43 \pm 3\ cm$, $c = 47 \pm 2\ cm$, so $\hat{D} = \hat{F} = \hat{a} * \hat{b} * \hat{c}$



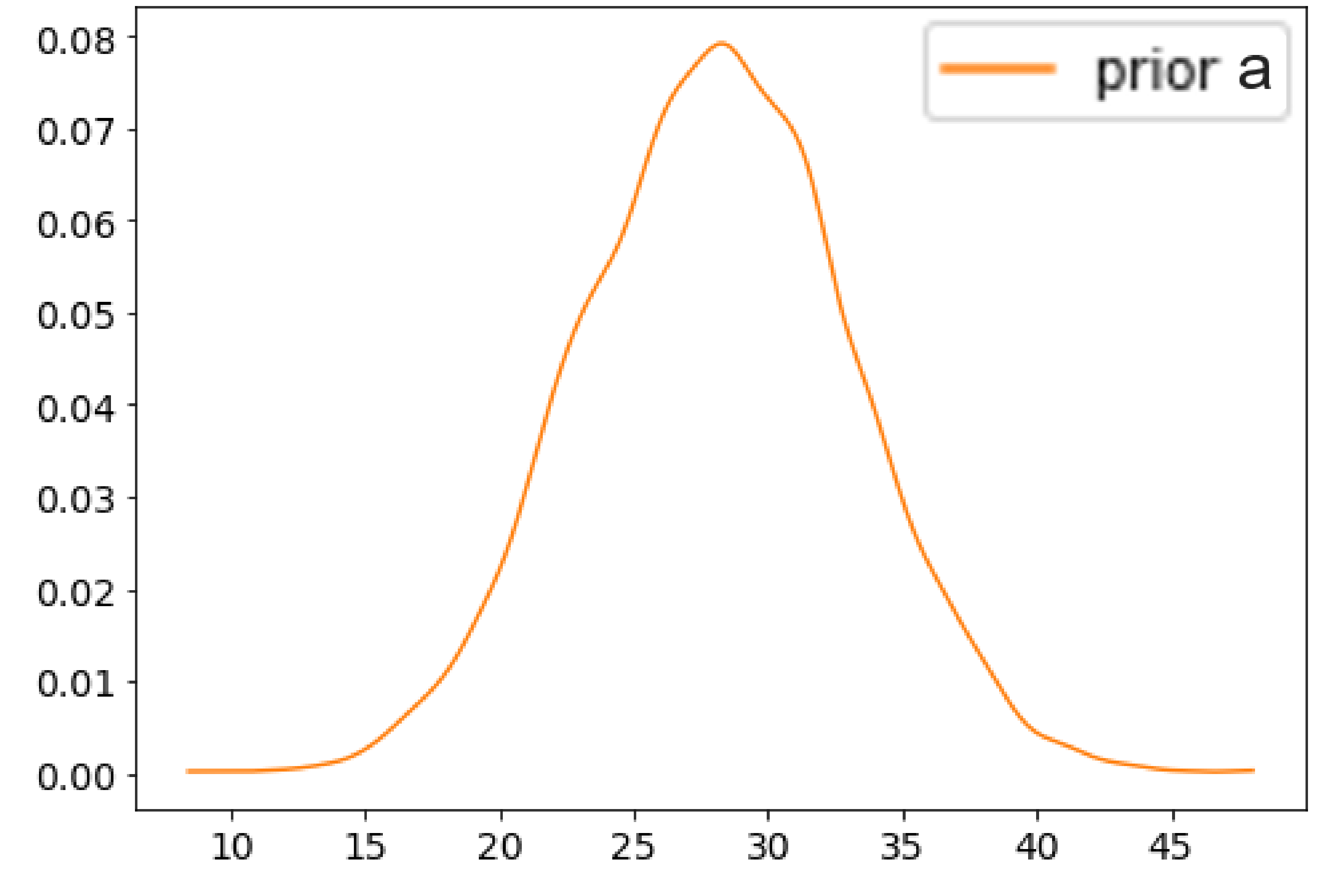
Sunnåker M, Busetto AG, Numminen E, Corander J, Foll M, et al. (2013) Approximate Bayesian Computation. PLoS Comput Biol 9(1): e1002803

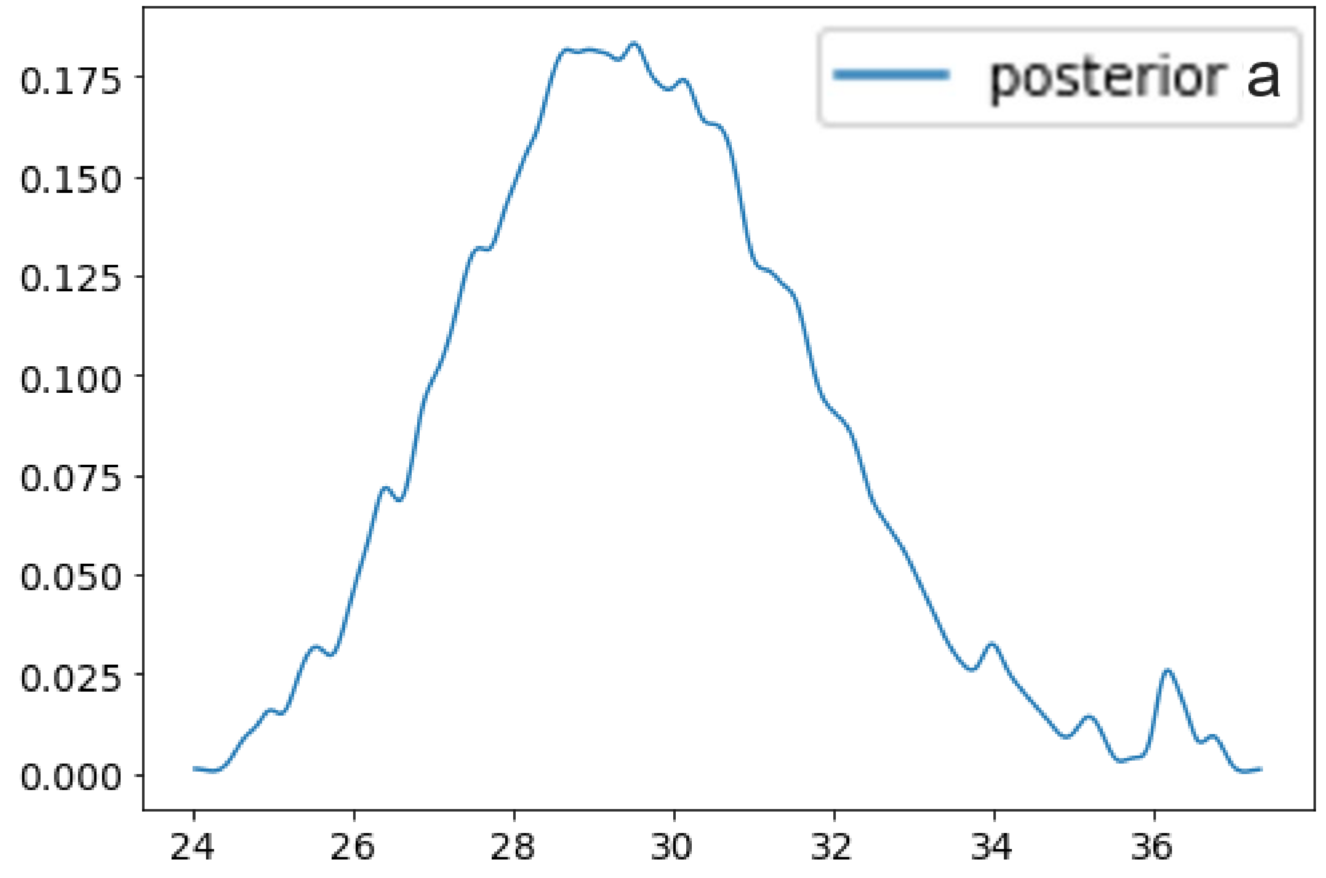
[1] Tom Leyshon - An introduction into parameter inference using Approximate Bayesian Computational methods, https://towardsdatascience.com/the-abcs-of-approximate-bayesian-computation-bfe11b8ca341

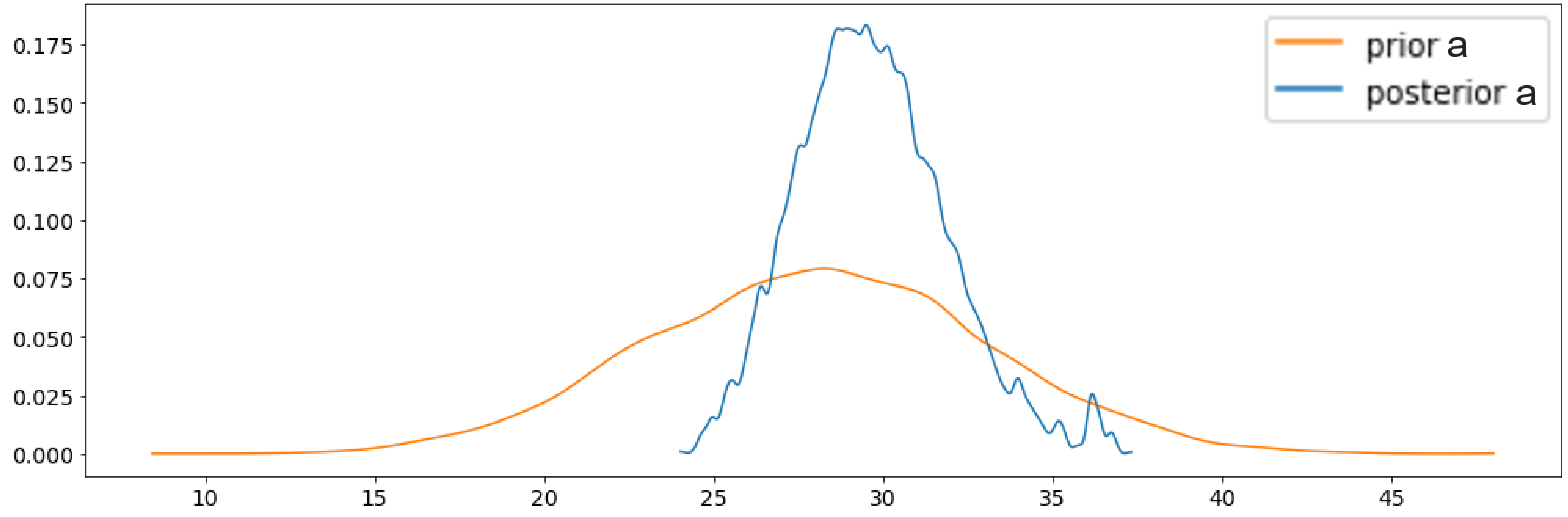
Calibration results

b

	a [cm]	b [cm]	c [cm]
prior	28 ± 5	43 ± 3	47 ± 2
posterior	29.6 ± 1.6	43.3 ± 2.5	47.01 ± 1.8
real value	30	40	50







F calculated from new a, b, c values: F = 60284 ± 5684 cm^3

a is closer to real value and uncertainties are reduced but b,c are further from real value – a sign of overfitting

What if we gather more data?

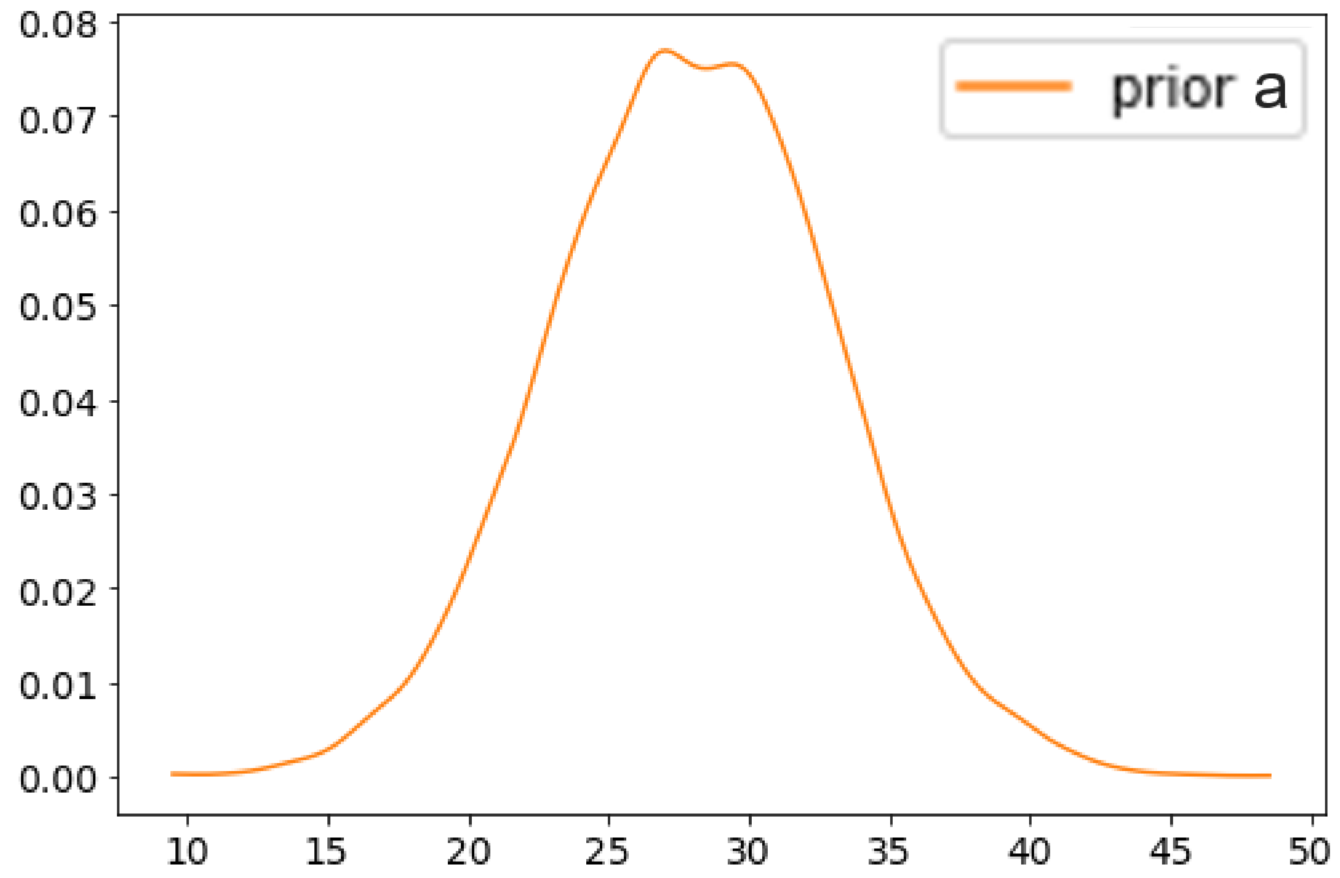
2 additional cuboids are constructed, with edges as in mathematical models below:

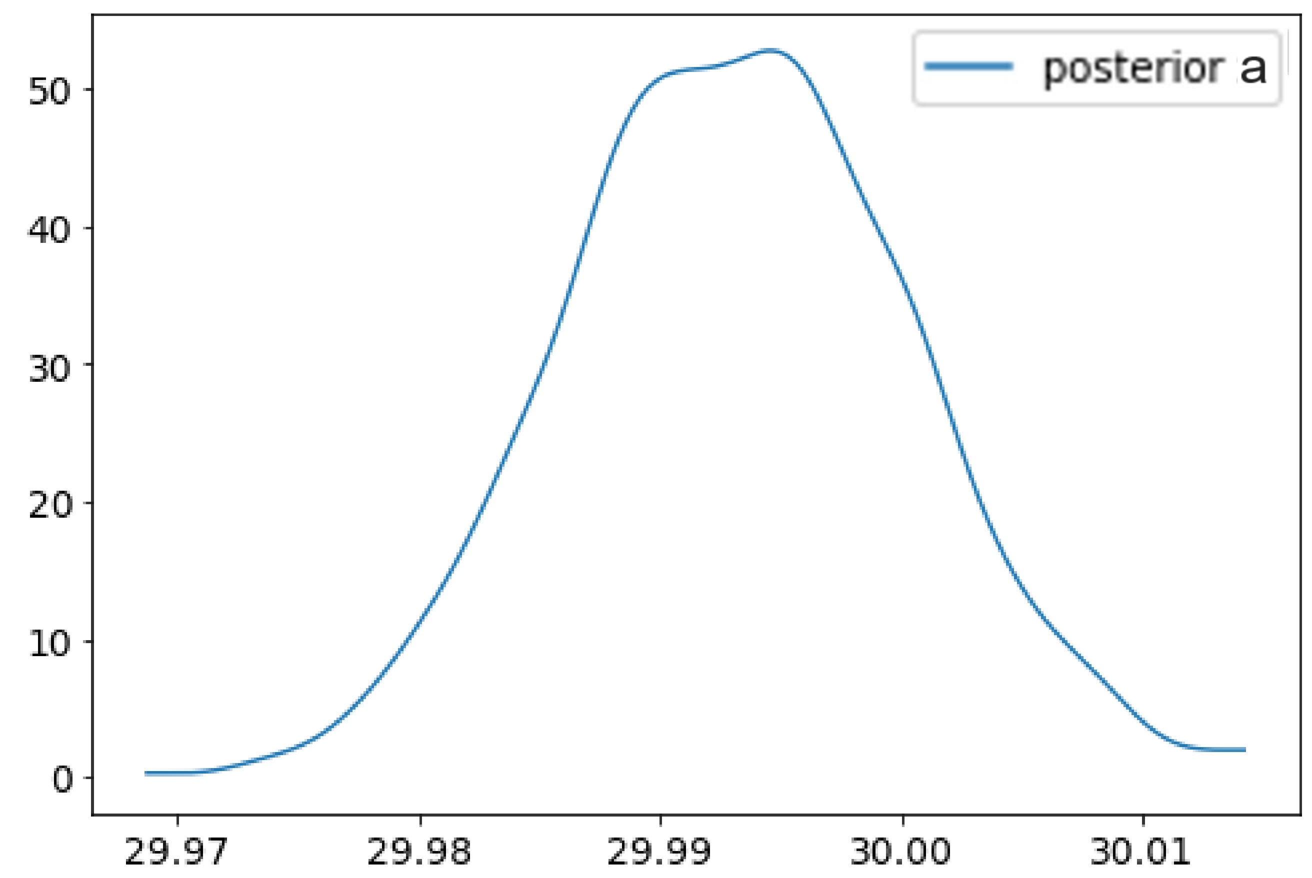
$$F_2 = a * a * b,\ F_3 = b * b * c$$

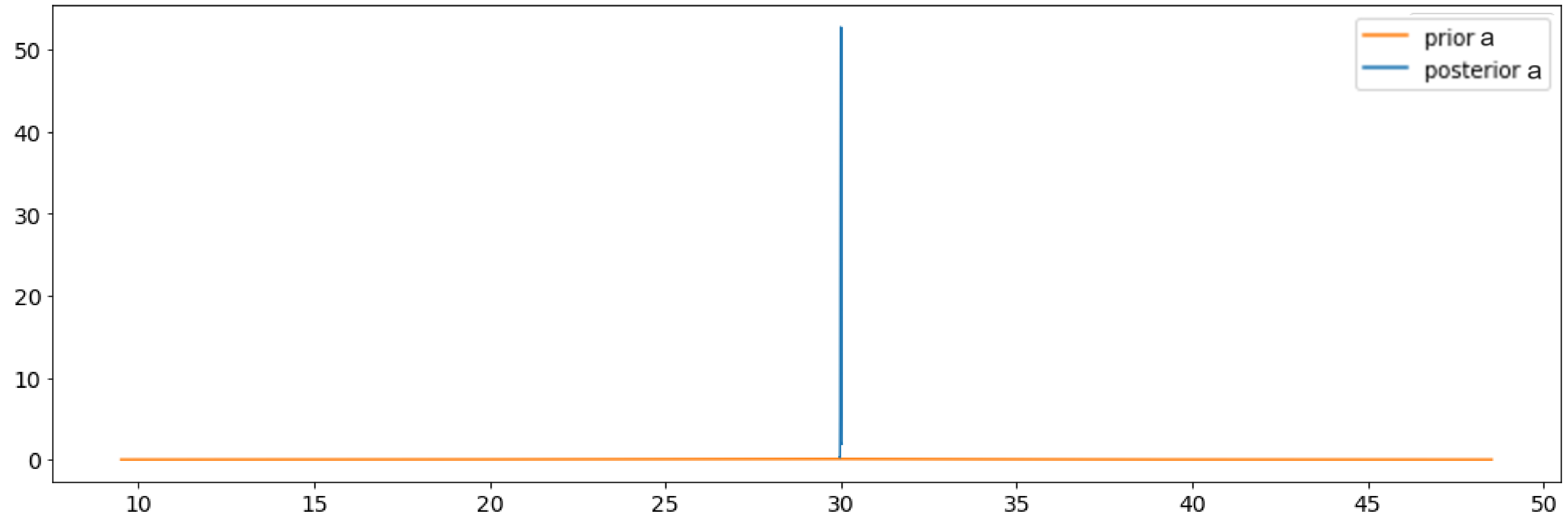
Cuboid number	Measured volume [cm³]	True volume [cm³]
1	60030 ± 20	60000
2	35975 ± 20	36000
3	80040 ± 20	80000

Results after including additional experiments

	a [cm]	b [cm]	c [cm]
prior	28 ± 5	43 ± 3	47 ± 2
posterior	29.993 ± 0.007	39.991 ± 0.013	50.047± 0.030
real value	30	40	50



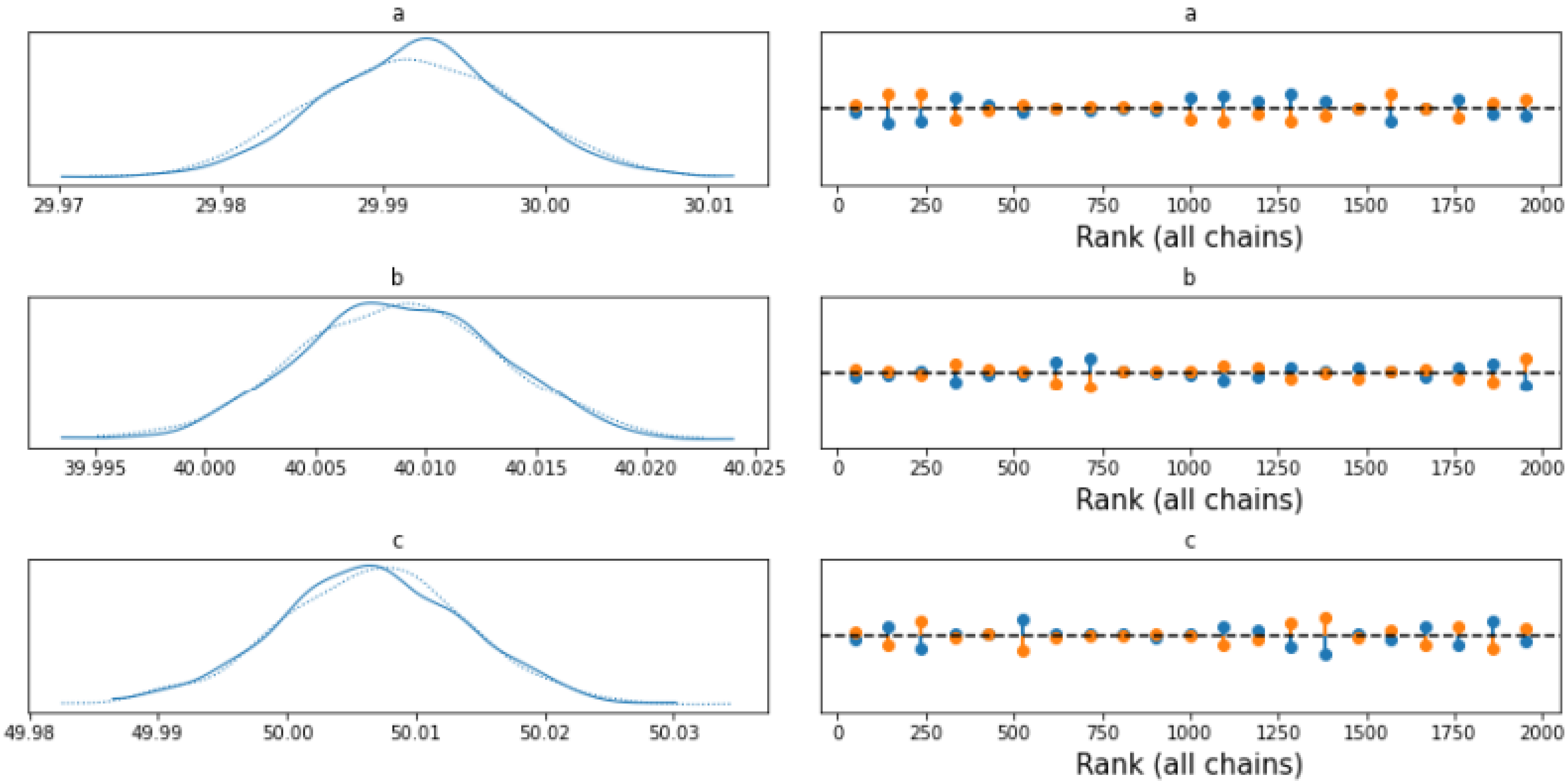




New parameter values are much closer to true values and have much lower uncertainty.

Superiority over other data fitting techniques: posterior uncertainty is given for each parameter.

Posterior distribution for each of rod lengths



Summary of components of ABC input

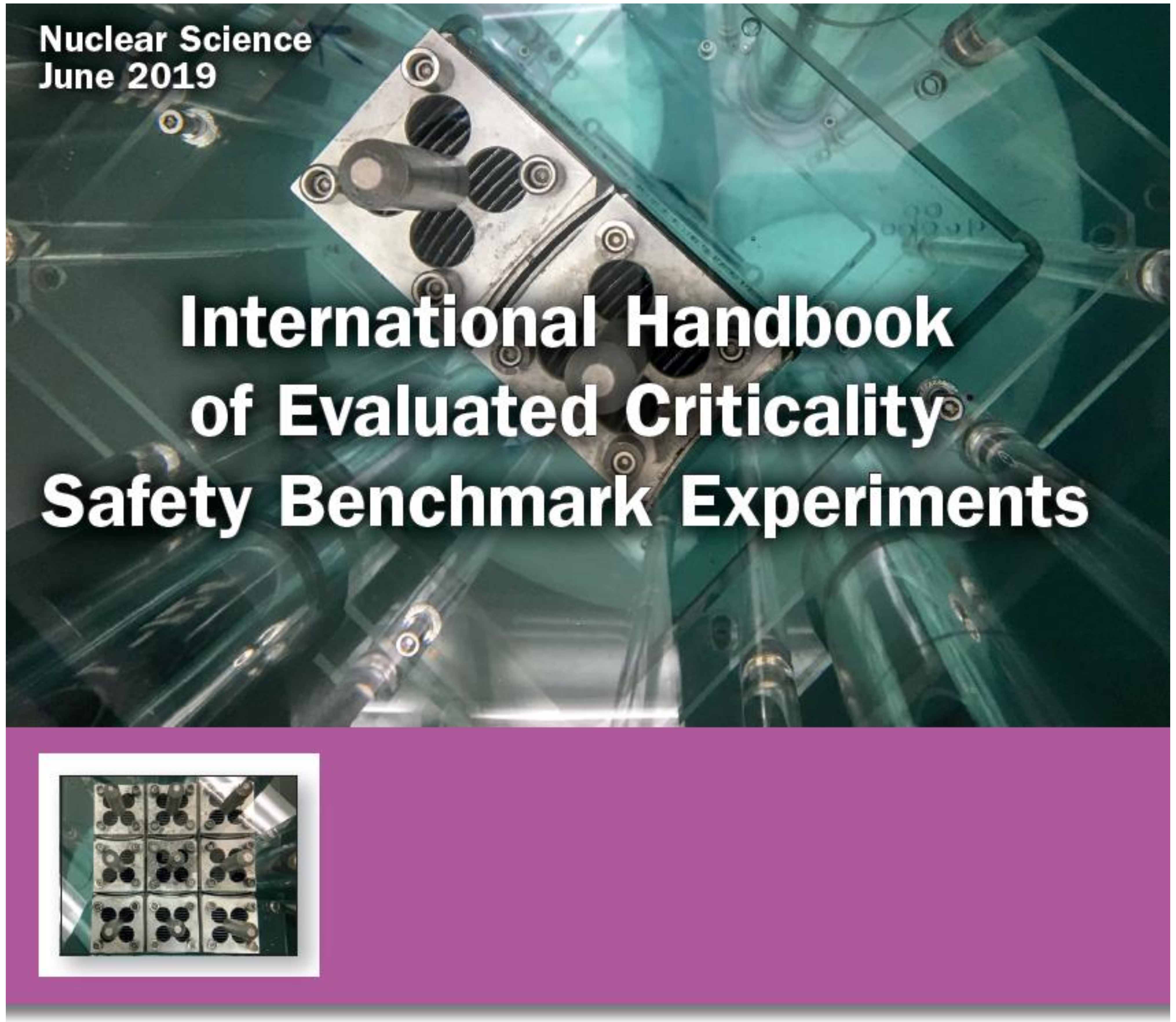
	Prior distribution	Mathematical model	Observed data	Uncertainty in observed data	Model discrepancy	Uncertainty from unadjusted parameters
Example above	$a = 28 \pm 5\ cm$ etc.	$F = a * b * c$, $F = a * a * c$, etc.	Cuboid volumes	20 cm^3	Omitted	All parameters are adjusted

Now, back to nuclear data

ABC input in my research

	Prior distribution	Mathematical model	Observed data	Uncertainty in observed data	Model discrepancy	Uncertainty from unadjusted parameters
Example above	$a = 28 \pm 5\ cm$ etc.	$F = a * b * c$, $F = a * a * c$, etc.	Cuboid volumes	20 cm^3	Omitted	All parameters are adjusted
Keff calculations	23 cross sections – energy averaged values + covariance matrix (multivariate normal distribution)	SCALE - nuclear simulation software	International Handbook of Evaluated Criticality Safety Benchmark Experiments – 24 experiments	Uncertainty due to idealizations and simplifications of geometry – usually 0.1%	Uncertainty due to using energy-averaged cross section values instead of continuous (0.4% of keff) + Monte Carlo uncertainty	Uncertainty from unadjusted cross sections

OECD/NEA experiment database



NEA
NUCLEAR ENERGY AGENCY

EVALUATED EXPERIMENTS

Highly Enriched Uranium Systems

METAL SYSTEMS

FAST METAL SYSTEMS

HEU-MET-FAST-001	Bare, Highly Enriched Uranium Sphere (Godiva)
HEU-MET-FAST-002	Topsy 8-Inch-Tuballoy-Reflected Oralloy Assemblies
HEU-MET-FAST-003	Reflected Oralloy Spherical Assemblies
HEU-MET-FAST-004	Water-Reflected, Highly Enriched Uranium Sphere
HEU-MET-FAST-005	Beryllium- and Molybdenum-Reflected Cylinders of Highly Enriched Uranium
HEU-MET-FAST-006	Lattices of Oralloy Cubes in Water
HEU-MET-FAST-007	Uranium Metal Slabs Moderated with Polyethylene, Plexiglas, and Teflon
HEU-MET-FAST-008	Bare Sphere of Highly Enriched Uranium
HEU-MET-FAST-009	Spheres of Highly Enriched Uranium Reflected by Beryllium or Beryllium Oxide
HEU-MET-FAST-010	Spheres of Highly Enriched Uranium Reflected by Boron+Beryllium or Boron+Beryllium Oxide
HEU-MET-FAST-011	Sphere of Highly Enriched Uranium Reflected by Polyethylene

Cross-sections and experiments chosen for the process

Cross-sections chosen for calibration:

(n, gamma)4
(n, gamma)6
(n, gamma)7
(n, gamma)9
(n, gamma)10
(n, gamma)11
(n, gamma)12
(n, gamma)13
(n, gamma)14
(n, gamma)15
(n, gamma)17
(n, n')4
(n, n')7
(n, n')9
(n,n')10
(n, n')11
(n, elastic)14
(n,fission)3
(n,fission)4
(n,fission)7
(n,fission)11
(n,fission)14
(n,chi)1

Experiments used in calibration:

HEU-MET-FAST-001-001
HEU-MET-FAST-007-004
HEU-MET-FAST-015-001
HEU-MET-FAST-016-001
HEU-MET-FAST-017-001
HEU-MET-FAST-018-001
HEU-MET-FAST-020-001
HEU-MET-FAST-021-001
HEU-MET-FAST-022-001
HEU-MET-FAST-025-001
HEU-MET-FAST-025-004
HEU-MET-FAST-032-001
HEU-MET-FAST-041-004
HEU-MET-FAST-065-001
HEU-MET-FAST-069-001
HEU-MET-FAST-084-001
HEU-MET-FAST-085-005
IEU-MET-FAST-001-004
IEU-MET-FAST-002-001
IEU-MET-FAST-003-001
IEU-MET-FAST-004-001
IEU-MET-FAST-005-001
IEU-MET-FAST-006-001
IEU-MET-FAST-009-001

My algorithm using SMC-ABC version of ABC

20000 cross-section sample sets are generated from prior distribution

↓

Each of 24 experiments is simulated with each sample set using SCALE (480000 SCALE inputs)

↓

The resulting keff values are gathered. ABC algorithm analyses them and generates new samples

↓

Each of 24 experiments is simulated with each sample set using SCALE (480000 SCALE inputs)

↓

(…)

↓

After some defined conditions are met, 20000 samples are chosen by algorithm to construct posterior distributions

Metamodel results

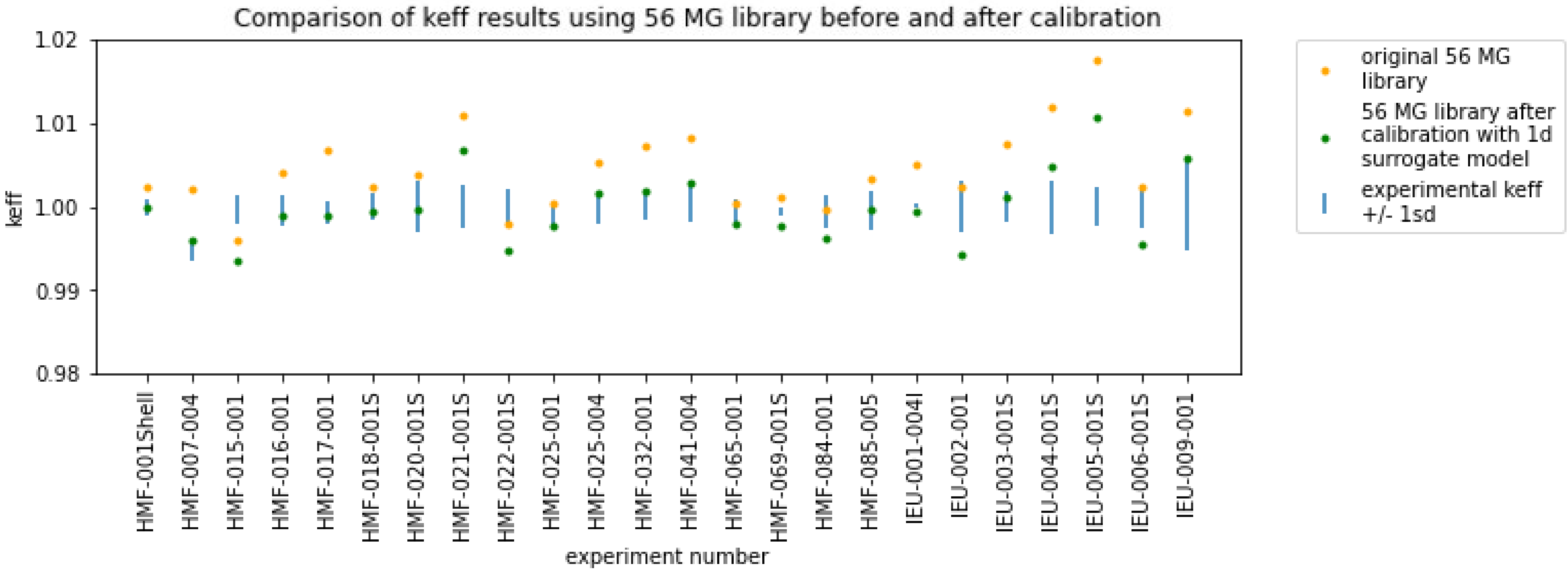
480000 SCALE computations per sample set is a lot. It is useful to first check if the proces has chance of succeding. This can be done using a coarser, less computationally demanding model.

SCALE TSUNAMI module generates sensitivities for each cross section, which can be used to create a 1D surrogate model (metamodel). Completing process using 1D metamodel takes 20 minutes on personal computer

Cross-section number	Value before ABC	Standard deviation before ABC	Value after ABC	Standard deviation after ABC
(n, gamma)4	1	0.220	0.980	0.208
(n, gamma)6	1	0.169	1.026	0.141
(n, gamma)7	1	0.160	1.049	0.115
(n, gamma)9	1	0.168	1.085	0.083
(n, gamma)10	1	0.180	1.105	0.070
(…)				
(n, fission)11	1	0.006	1.001	0.006
(n, fission)14	1	0.005	1.001	0.005
(n, chi)1	1	0.202	1.037	0.197

Validation of results

Comparison of keff results before and after using ABC generated cross-sections.



Average absolute difference between calculated and benchmark keff reduced by 49% (from 0.00549 to 0.00284)

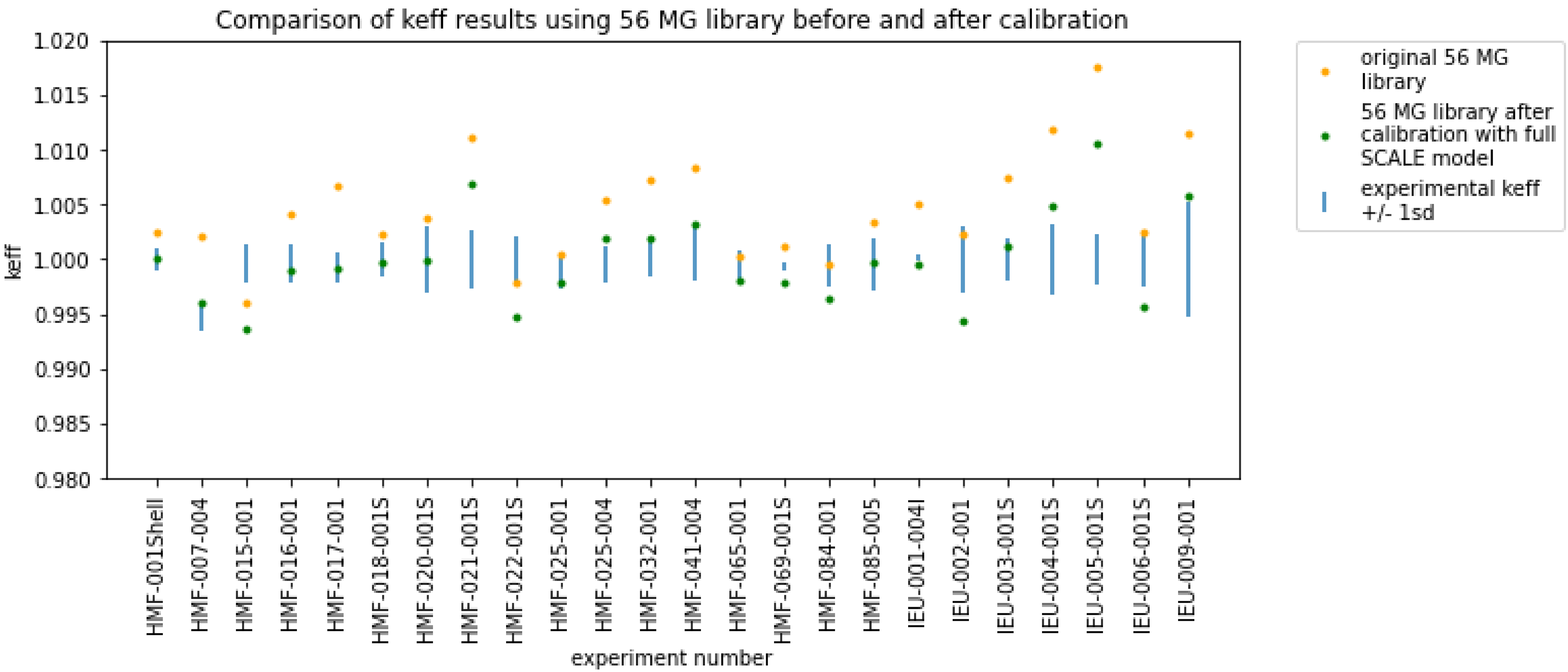
Calculations with SCALE as mathematical model

The process requires calculating keff for 20000 sample sets, 24 experiments each, 31 times. Completing the procedure once took 3 months using 4000 cores on NCBJ cluster.

New cross section values are identical to those generated by metamodel-driven ABC process.

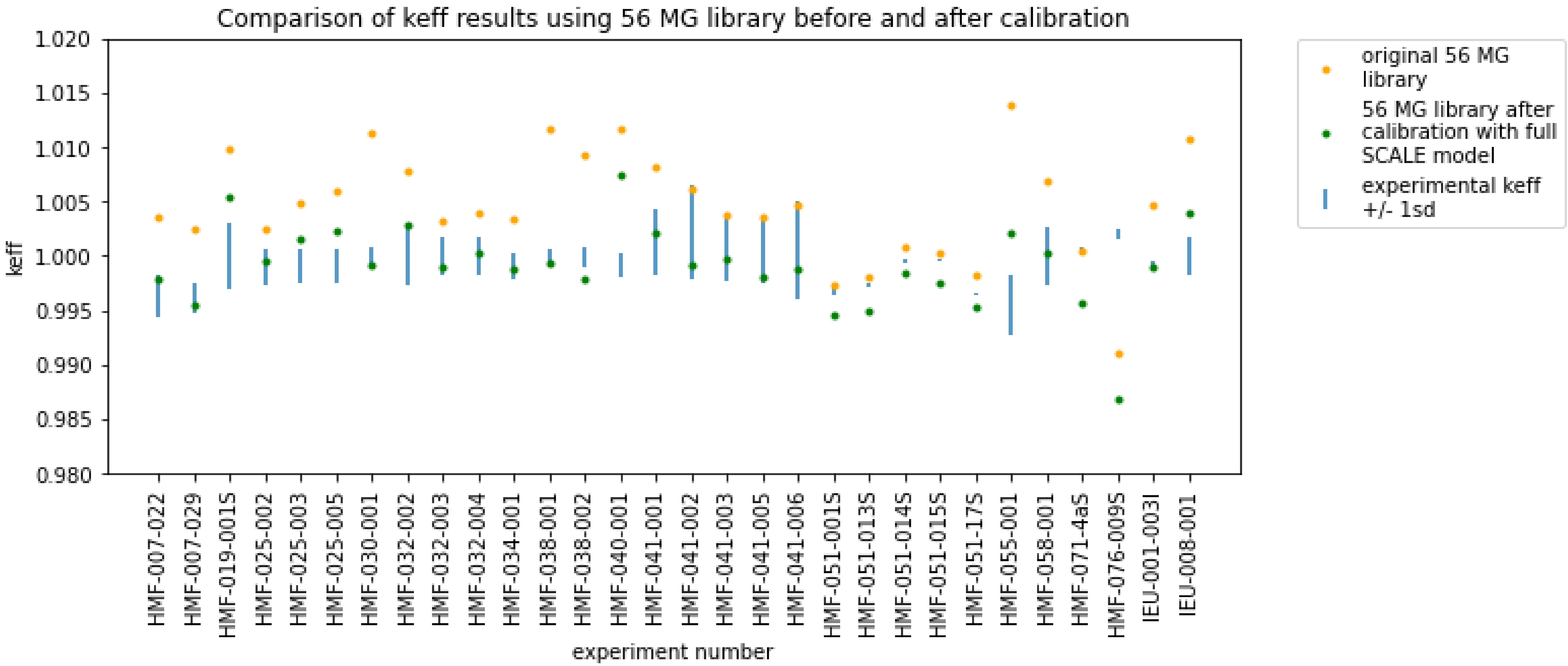
Cross-section number	Value after metamodel ABC	Standard deviation	Value after ABC SCALE model	Standard deviation
(n, gamma)4	0.980	0.208	0.971	0.216
(n, gamma)6	1.026	0.141	1.019	0.145
(n, gamma)7	1.049	0.115	1.042	0.118
(n, gamma)9	1.085	0.083	1.079	0.083
(n, gamma)10	1.105	0.070	1.099	0.069
(…)				
(n, fission)11	1.001	0.006	1.001	0.006
(n, fission)14	1.001	0.005	1.000	0.005
(n, chi)1	1.037	0.197	1.029	0.192

Validation of results using SCALE



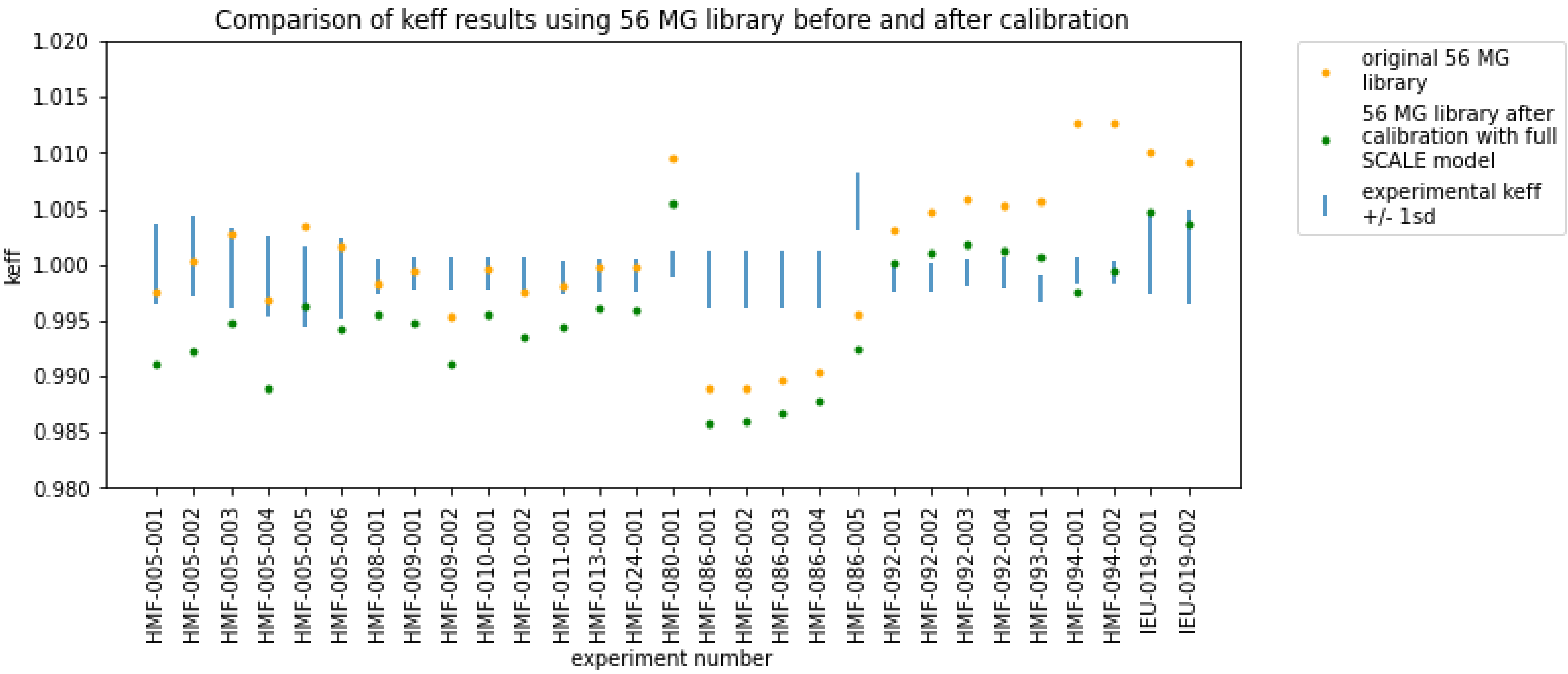
Average absolute difference between calculated and benchmark keff reduced by 49% (from 0.00549 to 0.00281)

Validation of results using SCALE



Average absolute difference between calculated and benchmark keff reduced by 56% (from 0.00591 to 0.00281)

Validation of results using SCALE

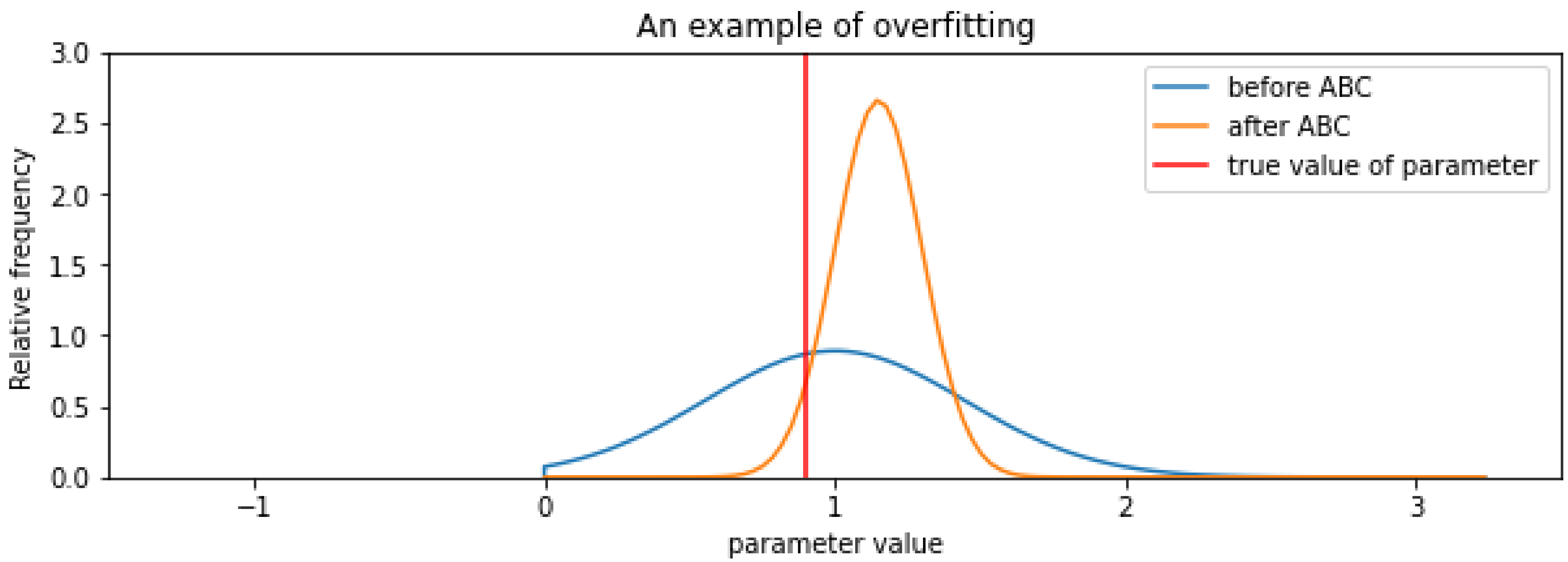


Average absolute difference between calculated and benchmark keff increased by 3% (from 0.00539 to 0.00555)

Reminder of the main risk when using ABC

What is the main risk when using ABC? – Overfitting, with possible causes as follows:

- too small number of experiments
- not diverse enough set of experiments
- too large uncertainty coming from uncalibrated parameters

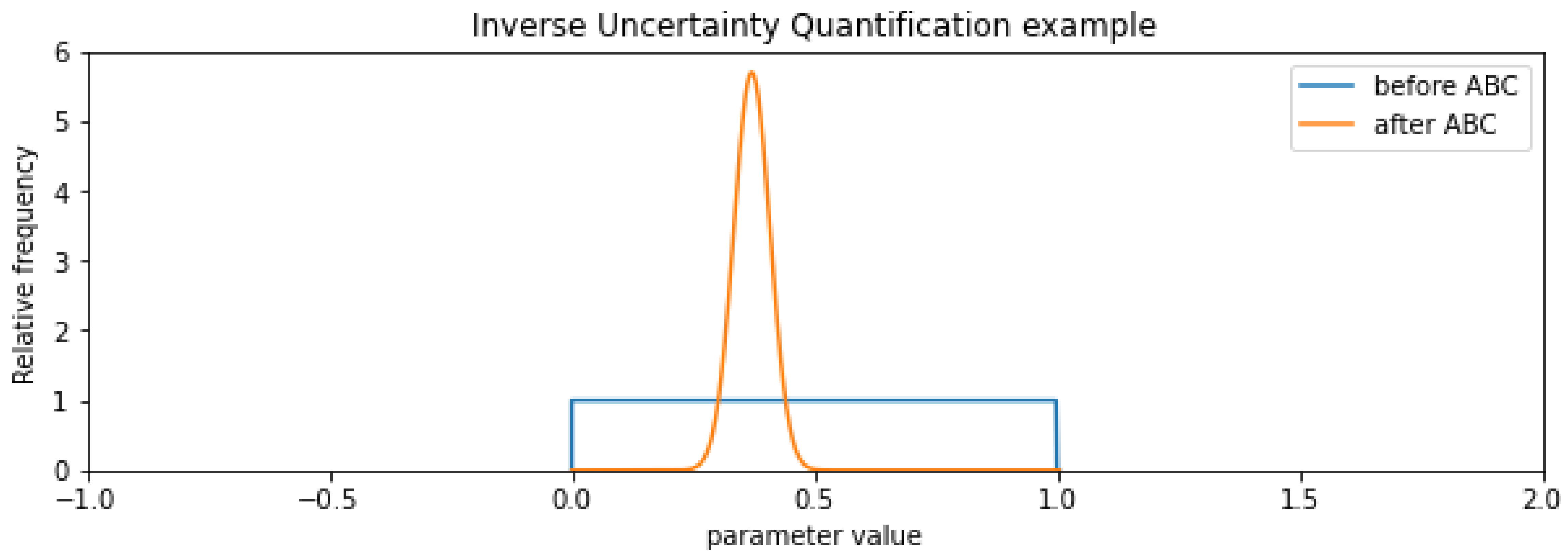


Conclusions and remarks

- ABC is a good method for reducing uncertainties of model parameters but re quires very high quality experiments and rigorous validation
- The more parameters we want to calibrate, the more computationally demanding the process is - calibrating 3 parameters may just require 1000 sample sets per draw, 23 parameters require at least 20000 sample sets – which means at least 20x more computation, likely more
- ABC using 1d metamodel can be done in 20 minutes on a personal computer, but more sophisticated model where each input requires minutes for computation may take months to finish on cluster

Another possible use of ABC

Assesing uncertainty of dubious model parameters, like empirical parameters in thermohydraulic problems.

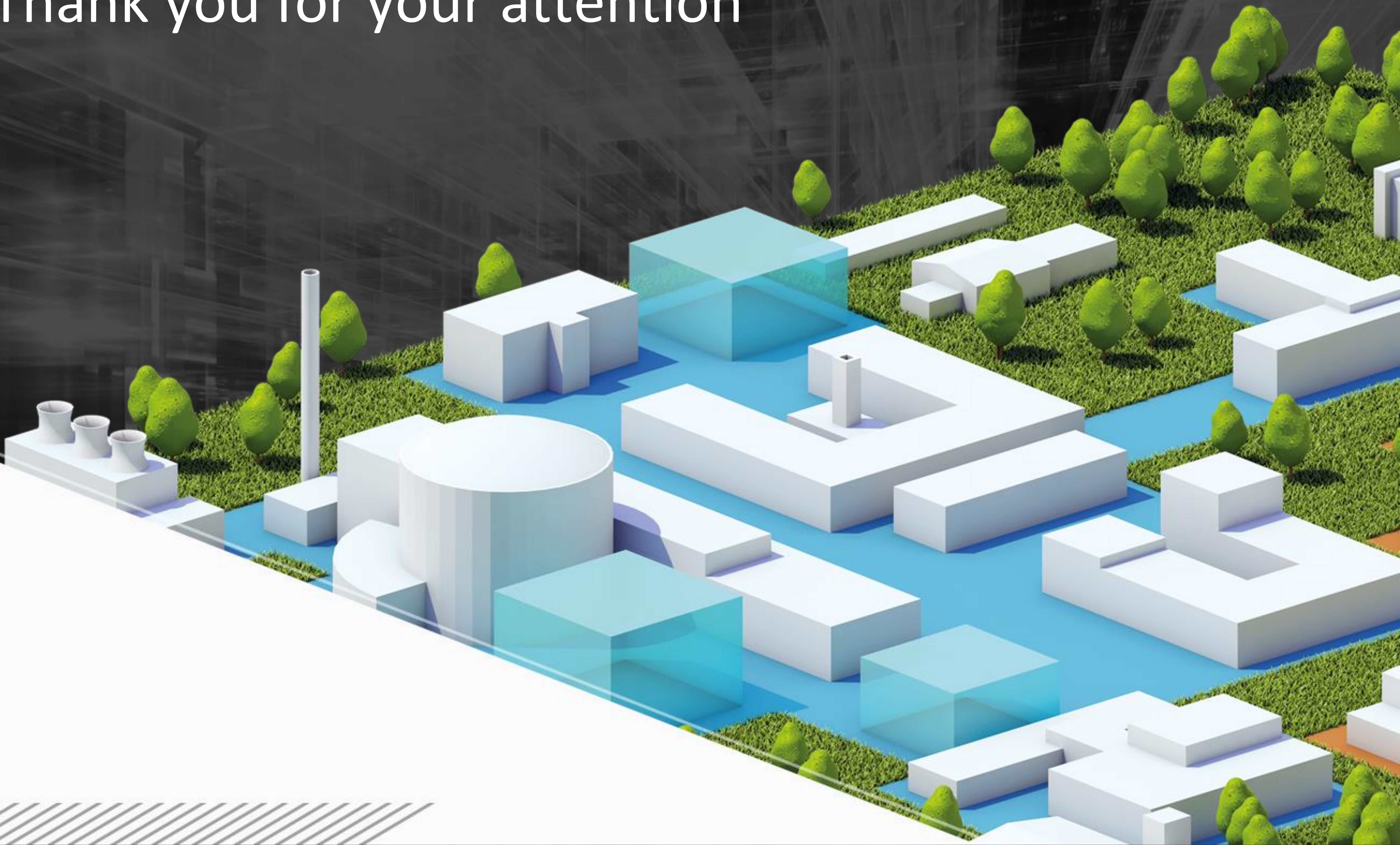


• X. Wu, T. Kozlowski et al., Inverse uncertainty quantification using the modular Bayesian approach based on Gaussian process, Part 2: Application to TRACE, 2018

• Ziyu Xiea, Wen Jiang et al., Bayesian Inverse Uncertainty Quantification of a MOOSE-based Melt Pool Model for Additive Manufacturing Using Experimental Data, 2021

• X. Wu, T. Kozlowski, H. Meidani., Kriging-based inverse uncertainty quantification of nuclear fuel performance code BISON fission gas release model using time series measurement data, 2018

Thank you for your attention



NATIONAL CENTRE FOR NUCLEAR RESEARCH ŚWIERK

www.ncbj.gov.pl