GRS

Generation of Integral
Experiment Covariance
Data and their Impact
on Criticality Safety
Validation

GRS

Generation of Integral Experiment Covariance Data and their Impact on Criticality Safety Validation

Maik Stuke
Elisabeth Peters
Fabian Sommer

November 2016

Remark:

Kurzfassung

Mögliche Quantifizierungen statistischer Abhängigkeiten in Daten kritischer Experimente und dessen Berücksichtigung bei Validierungsverfahren, ist in der Literatur von verschiedenen Gruppen erörtert worden und ist zur Zeit eine aktiv geführte Diskussion in der Expertengruppe zur Unsicherheitsanalyse für die Sicherheitsprüfung *Expert Group on Uncertainty Analysis for Criticality Safety Assessment* (UACSA) des OECD- NEA-Ausschuss für Nuklearwissenschaften, dem *Nuclear Science Committee*. Letzterer erstellt und publiziert die frei verfügbaren experimentellen Daten im *International Handbook of Evaluated Criticality Safety Benchmark Experiments*, ICSBEP. Die meisten Experimente wurden als Serie durchgeführt und teilen sich einige der experimentellen Komponenten, was zu Korrelationseffekten in den Ergebnissen führen kann. Die korrekte Betrachtung korrelierter Daten scheint unvermeidbar zu sein, wenn die Anzahl der Experimente in einem Validierungsverfahren begrenzt ist oder man sich nicht auf eine ausreichende Anzahl von unkorrelierten Datensätzen stützen kann, z.B. Daten verschiedener Labors mit verschiedenen Komponenten. Die generelle Bestimmung der Korrelationen und der zugrunde liegenden Kovarianzdaten sowie deren Berücksichtigung in einem Validierungsverfahren liegt im Fokus der folgenden Arbeiten.

Wir diskutieren und demonstrieren mögliche Effekte auf berechnete k_{eff}-Werte, deren Unsicherheiten und entsprechende Kovarianzmatrizen aufgrund der Interpretation ausgewerteter experimenteller Daten und deren Übersetzung in Berechnungsmodelle. Die Arbeit zeigt Effekte verschiedener Modellierungsansätze, variierender Verteilungsfunktionen von Parametern und vergleicht Ergebnisse der angewandten Monte-Carlo Stichprobenmethode mit verfügbaren Daten zu experimentellen Korrelationen. Unsere Ergebnisse zeigen, dass für die zuverlässige Bestimmung von integralen experimentellen Kovarianzmatrizen oder der Korrelationskoeffizienten eine detaillierte Untersuchung der zugrunde liegenden experimentellen Daten, des Modellierungsansatzes und der getroffenen Annahmen sowie der daraus resultierenden Sensitivitätsanalyse unvermeidbar erscheint.

Weiterhin wird ein Bayes-Verfahren diskutiert, um integrale experimentelle Kovarianzdaten bei der Schätzung des k_{eff} eines Anwendungsfalls einzuschließen. Es wird gezeigt, wie der berechnete k_{eff}-Wert eines Anwendungsfalles von den Kovarianzen abhängt und was die Auswirkungen von unberücksichtigten Korrelationen sein können.

Abstract

The quantification of statistical dependencies in data of critical experiments and how to account for them properly in validation procedures has been discussed in the literature by various groups. However, these subjects are still an active topic in the *Expert Group on Uncertainty Analysis for Criticality Safety Assessment* (UACSA) of the OECD-NEA Nuclear Science Committee. The latter compiles and publishes the freely available experimental data collection, the *International Handbook of Evaluated Criticality Safety Benchmark Experiments*, ICSBEP. Most of the experiments were performed as series and share parts of experimental setups, consequently leading to correlation effects in the results. The correct consideration of correlated data seems to be inevitable if the experimental data in a validation procedure is limited or one cannot rely on a sufficient number of uncorrelated data sets, e.g. from different laboratories using different setups. The general determination of correlations and the underlying covariance data as well as the consideration of them in a validation procedure is the focus of the following work.

We discuss and demonstrate possible effects on calculated k_{eff}'s, their uncertainties, and the corresponding covariance matrices due to interpretation of evaluated experimental data and its translation into calculation models. The work shows effects of various modeling approaches, varying distribution functions of parameters and compares and discusses results from the applied Monte-Carlo sampling method with available data on correlations. Our findings indicate that for the reliable determination of integral experimental covariance matrices or the correlation coefficients a detailed study of the underlying experimental data, the modeling approach and assumptions made, and the resulting sensitivity analysis seems to be inevitable.

Further, a Bayesian method is discussed to include integral experimental covariance data when estimating an application case k_{eff}. It is shown, how the calculated k_{eff} of an application case depends on the covariances and what would be the impact of ignoring them.

Content

1   Introduction

Criticality safety assessments require a prediction of the effective neutron multiplication factor k_{eff} below a sufficient safety margin. This predicted value is derived using a validated calculation method with validated computer codes, e.g. so called criticality codes to calculate the k_{eff} of a given application case. The validation of a criticality code can be achieved by recalculations of suitable critical experiments performed in laboratories and documented and evaluated e.g. in [1]. In recent years, several authors discussed the fact that depending on the application case and the choice of experiments, the effect of correlated experimental data on the determination of the bias, its uncertainty, and the resulting safety margins has to be considered [2, 3, 4, 5, 6, 7, 8, 9, 10]. The questions arising in the field of determination and handling of integral experimental covariance matrices in the process of code validation are also discussed in the Expert Group on Uncertainty Analysis for Criticality Safety Assessment (UACSA), a sub-group of the Working Party on Nuclear Criticality Safety (WPNCS) of the Nuclear Energy Agency (NEA) within the Organization for Economic Co-operation and Development (OECD). Some of the current questions are: How to treat given sets of similar experimental data without knowing all exact statistical dependencies; and further, what are the implications on modeling these experiments in a code validation procedure regarding the consideration of the complete integral experimental correlation or covariance matrices? The discussions of the UACSA are synthesized in the benchmark proposal [11].

Correlated data can arise if different experiments share parts of the experimental setup, measurement systems, or other relevant parameters. Some experiments described in the ICSBEP are not performed as single experiments, but slight variations of a setup were repeatedly investigated and published as a series of the same experiment. The *traditional* validation procedures do not consider any correlations, but treat every experiment as statistical independent. If one can ensure this independence, e.g. by considering only experiments from different laboratories with different nuclear fuel characteristics, the traditional methods are just fine. However, if the data base of suitable experiments in a validation procedure is limited and correlation effects can not be excluded, the ignoring of these correlations has the potential of generating non-conservative results. An important role in the determination of correlations has the translation of the experimental data into calculation models. Recent discussions in the UACSA have shown, that different groups derived different correlations even though the underlying experimental data was identical.

In the following we introduce the general statistic terms used in this work in chapter 2. In chapter 3 we discuss how we generated integral experiment covariance data and discuss and illustrate the main obstacles in this process. Chapter 4 discusses the impact of the generated covariance data on some application cases. We conclude our findings in section 5.

2   General Statistics

The following is intended to give a brief overview of statistics and to introduce the used statistical terms and definitions. More details can be found in any of the numerous available textbooks on statistics. A good and compact overview can be found e.g. in [12, 13]. Concerning the mathematical notation used in this manuscript we tried to remain as close as possible to already published notations on the same subject. This holds especially for the notation of the Bayesian updating process, introduced in [7].

Frequently, one is confronted with an accumulation of data and the wish to use this data to derive and gather general statements from it. This data could be e.g. a set of experimental observations or samples of theoretical predictions. Statistics can be used to derive inference about probabilistic models, e.g. the validity of a model or certain parameters it predicts. Supposing the experimental data consists of independent measurements $(x_1, ..., x_n)$ and its parameters following a certain probability density function (p.d.f.), the unbiased estimators of the principally unknown mean value μ and variance σ^2 are

$$\hat{\mu} \;=\; \frac{1}{n} \sum_{i=1}^{n} x_i, \tag{2.1}$$

$$\hat{\sigma}^2 \;=\; \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \hat{\mu})^2 . \tag{2.2}$$

The variance of $\hat{\mu}$ is σ^2/n and the variance for Gaussian distributed x_i is $2\sigma^4/(n-1)$ for $n \geq 2$. For large n, the *error of the error* becomes $\sigma/\sqrt{2n}$.

For any n and Gaussian distributed x_i, $\hat{\mu}$ is an efficient estimator for μ, and $\hat{\mu}$ and $\hat{\sigma}$ are uncorrelated. Note, that following equation 2.2, knowing σ^2 does not improve the estimator $\hat{\mu}$. But knowing μ does improve the variance of the estimator of σ^2.

If the x_i have different but known variances, the weighted average

$$\hat{\mu} = \frac{1}{w} \sum_{i=1}^{n} w_i x_i, \tag{2.3}$$

with $w_i = 1/\sigma_i$ and $w = \sum w_i$, is an unbiased estimator for μ with a smaller variance than the unweighted average. The standard deviation of $\hat{\mu}$ becomes $1/\sqrt{w}$.

Dealing with uncertainties or errors of experimental values, it can become necessary to

assign a distribution function to errors to be able to apply certain statistical tools, e.g. Monte-Carlo sampling or Bayesian statistics. For the purpose of this work it is a very good approximation to assume all errors following a normal distribution. The reason is based on the central limit theorem. It basically states, that as the number of variables in a sum increases the distribution of the sum of random variables approaches the normal distribution regardless of the shape of the distribution of the individual random variables. In general, an experimental error consists of a large number of contributing errors. If the sources of errors are numerous then by the central limit theorem one can say that experimental errors tend to have a normal distribution. Using statistical tools to make inference about a sample average, which is a further summation of values that are themselves likely to have normally distributed errors, provides additional justification to assume that sample averages have normal distributed errors. However, it is worth noting, that the errors of observations generally tend to be normal distributed, but not the random variable itself. The latter can follow any distribution function, depending e.g. on physical properties or definitions e.g. positive definite, or following a uniform distribution.

To compare two random variables or to investigate the dependence between them one usually calculates the covariance. The covariance of two random variables is a measure of how much the two variables change together. It has a positive value, if greater (lesser) values of one variable corresponds to greater (lesser) of the other. If greater (lesser) values of one variable correspond to lesser (greater) of the other, the covariance is negative. For two sets A and B of n sampled neutron multiplication factors k_{eff}, the covariance cov^{AB} is defined as

$$\text{cov}_{\text{AB}} = \frac{1}{n-1} \sum_{i=1}^{n} \left(k_{\text{eff}}^{\text{A},i} - \overline{k_{\text{eff}}}^{\text{A},i} \right) \left(k_{\text{eff}}^{\text{B},i} - \overline{k_{\text{eff}}}^{\text{B},i} \right) \tag{2.4}$$

with $\overline{k_{\text{eff}}}$ symbolizing the expectation value of k_{eff}, here the sample mean.

To get comparable statements for more than two sets of random variables the covariance can be normalized with the standard deviation σ to get Pearson's correlation coefficient[1] cor:

$$\text{cor}_{\text{AB}} = \frac{1}{\sigma_A \sigma_B} \text{cov}_{\text{AB}}. \tag{2.5}$$

---

[1] In the following the expression *correlation coefficients* refers always to Pearson's definition.

The correlation coefficient is a non-additive and dimensionless measure of the linear dependence of two sets of random variables and takes values between +1 (complete positive linear connection) and -1 (complete negative linear connection). Fig. 2.1 and 2.2 show two different ways of visualizing correlation graphically for the same data set. Both picture show the correlation of nine different data sets named BM1 to BM9 consisting each of 250 entries. Both pictures show a matrix of these data sets, mirrored at the diagonal ($\mathrm{cor}_{AB} = \mathrm{cor}_{BA}$). Fig. 2.1 shows a matrix of scatter plots with the distribution of values for each of the nine data sets on the diagonal. The more the point clouds are linearly distributed the higher is the correlation value. The more circled distributed they are, the lower is the correlation value. This type of representation offers a lot of information apart from the correlation coefficient. E.g. the check of the distribution function on the diagonal axis, or to exclude correlations different than linear.

Fig. 2.2 shows the correlation coefficients as numbers in the upper right half and as color coded squares in the lower half. The diagonal elements show the color coded correlation coefficients $\mathrm{cor}_{AA} = 1$. The information provided in these pictures is less compared to fig. 2.1, but might appear more intuitive and readable. For this reason, we use in this publication only these types of visualizations.

Transforming the correlation coefficients via Fisher's z-transformation [14] gives an additive measure which can be transformed back. The z-distribution is shown in fig. 2.3 and is defined as

$$z(\mathrm{cor}) := 0.5 \ln \left(\frac{1 + \mathrm{cor}}{1 - \mathrm{cor}} \right). \tag{2.6}$$

The z-distribution is approximately normal distributed with a standard deviation

$$\sigma(z) = \sqrt{(n-3)^{-1}} \tag{2.7}$$

with n being the number of samples. Following this definition, one can compute a confidence interval CI of a correlation coefficient in z-space following

$$\mathrm{CI(z)} = \mathrm{z(cor)} \pm \mathrm{CL} \times \sigma(\mathrm{z}) \tag{2.8}$$

where CL is the confidence level. For the uni-variant case of a normal distributed random variable with known standard deviation, the probability that a value x falls within the
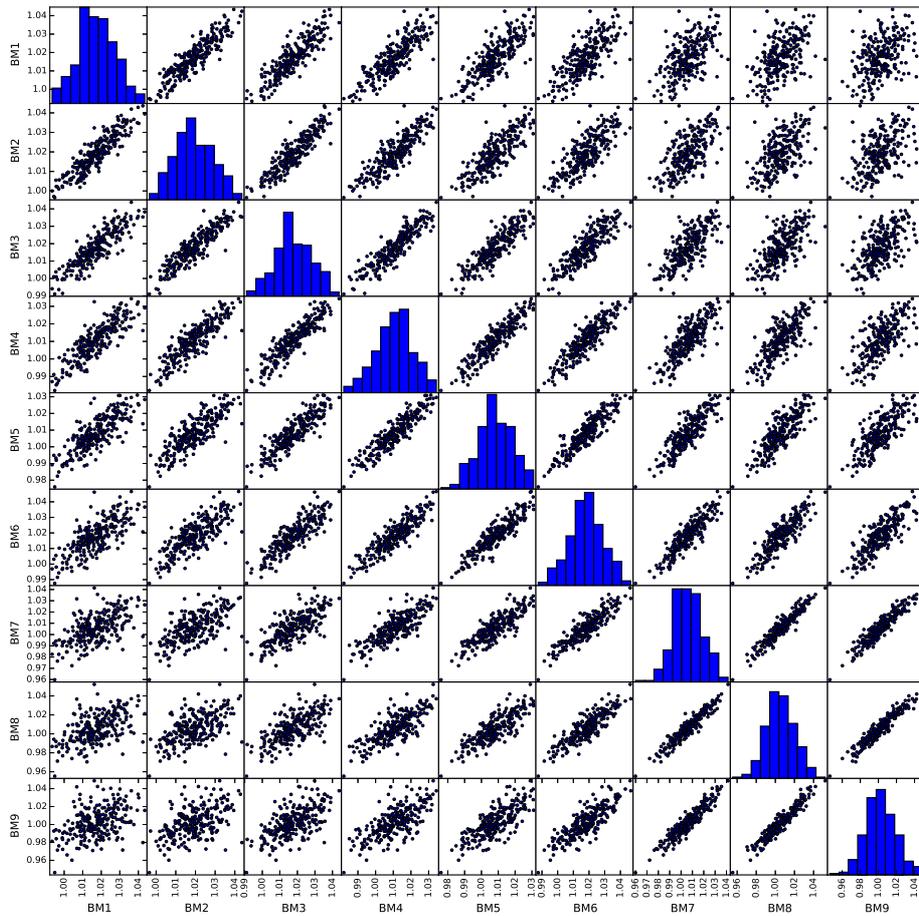
Fig. 2.1   Shown here is a matrix of correlation of nine different data sets consisting each of 250 entries. The matrices are mirrored at the diagonal and show a matrix of scatter plots with the distribution of the values for each of the nine data sets on the diagonal. The more the point clouds are linearly distributed the higher is the correlation value. The more circled distributed they are, the lower is the correlation value.
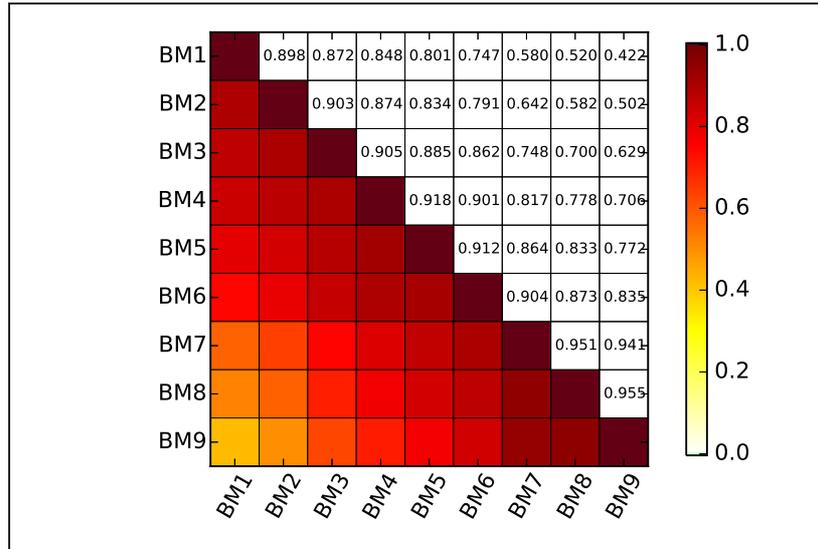
Fig. 2.2 Shown are the correlation coefficients as numbers in the upper right half and as color coded squares in the lower half of the matrix. The diagonal elements are per definition $\mathrm{cor}_{AA} = 1$.

interval $\pm\delta = \pm m\sigma$, with $m \in \Re$ around the true value μ is defined as

$$1 - \alpha \;=\; \frac{1}{\sqrt{2\pi}\sigma} \int_{\mu-\delta}^{\mu+\delta} \exp\left[-\frac{(\mathrm{x}-\mu)^2}{2\sigma^2}\right] \mathrm{dx} \tag{2.9}$$

$$\;=\; \mathrm{erf}\left(\frac{\delta}{\sqrt{2}\delta}\right) \tag{2.10}$$

with erf being the Gaussian error function. The choice $\delta = \sigma$ gives the interval called *standard error* with a confidence level of $1 - \alpha = 68.27\%$. Further values of corresponding pairs of standard deviations and confidence levels are e.g. $m = 1.64$ (90%), 1.96 (95%), 2.58 (99%), and 3.29 (99.9%). An illustration of the confidence interval is shown in fig. 2.4. The calculated upper and lower limits in z-space can be transformed back via

$$\mathrm{cor(z)} = \frac{\exp^{2\mathrm{z}} - 1}{\exp^{2\mathrm{z}} + 1} \tag{2.11}$$

to get the corresponding confidence interval for the calculated correlation coefficient.

2.1 Bayesian Statistic

In the following we will introduce and describe the relevant concepts and formulas. We will not go into too many details since there exists a vast number of publications on the subject of Bayes statistic and its advantages and disadvantages compared to the traditional or
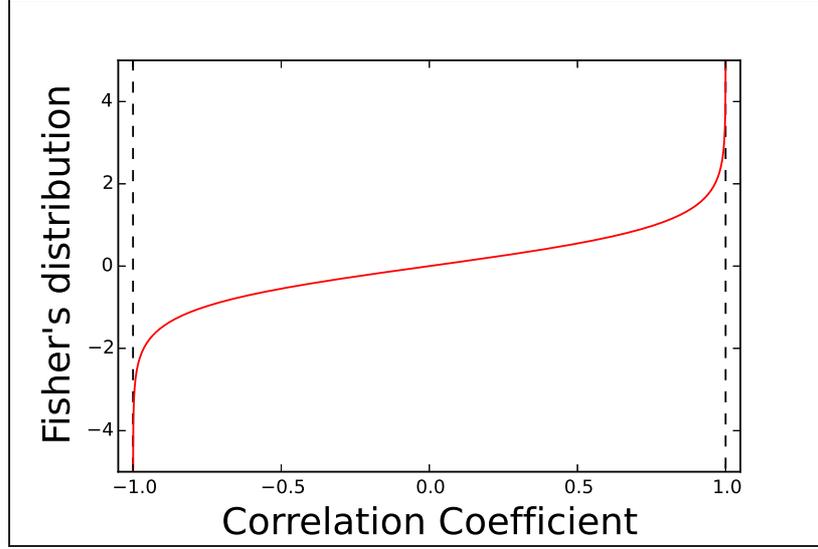
Fig. 2.3   Fisher z-distribution as a function of correlation coefficient. The dashed lines show the asymptotes at ± 1.

frequentist approach. In the context of criticality safety, applications of Bayesian Statistics is described e.g. in [15, 16].

In the concept of Bayes statistics, a prior believe of an event is updated with further knowledge. Technically it means to associate to an event A the probability to occur $p(A)$ and then gain a better knowledge by update it with further, related knowledge X as constrained probabilities $p(A|X)$, meaning event A occurs under the assumption of X. The Bayes-Theorem for the so called posterior probability density $P(A|X)$ is defined as

$$P(A|X) = (P(X|A)P(A))/(P(X)). \tag{2.12}$$

Expressed in likelihood functions it becomes $P(A|X) \propto L(A|B)p(A)$. Assuming independent realizations of a normal distributed variable $y = \theta + N(0,1)$ with unknown mean μ and variance σ^2, the Likelihood function is given as

$$L(y|\theta) = [y|\theta] = 1/\sqrt{\pi}\exp[-1/2(y-\theta)^2]. \tag{2.13}$$

With a Gaussian prior for $[\theta] = 1/(\sqrt{\pi}\sigma)\exp[-1/2(\theta-\mu)^2]$ we get for the posterior probability density

$$\begin{aligned}
[\theta|y] &\approx [y|\theta][\theta] && \text{(2.14)} \\
&= 1/\sqrt{\pi}\exp[-1/2(y-\theta)^2]1/(\sqrt{\pi}\sigma)\exp[-1/2(\theta-\mu)^2] && \text{(2.15)} \\
&\propto \exp[1/(\sqrt{\pi}\sigma^*)]\exp[-1/2(y^*-\theta)^2]. && \text{(2.16)}
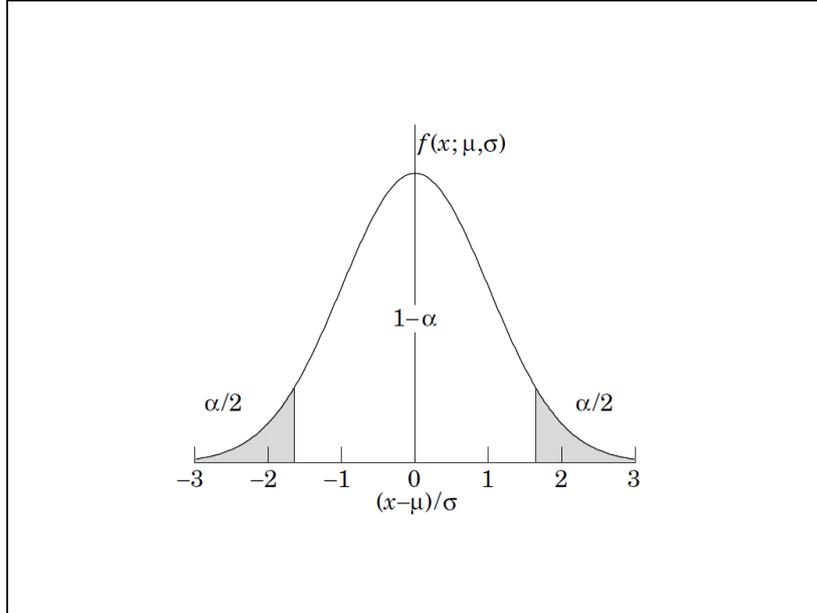\end{aligned}$$

8

Fig. 2.4 Illustration of a symmetric 90% confidence interval for a uni-variant Gaussian distribution.

The latter relation arises from the fact, that two Gaussian distributions result again in a gaussian. The asterisk marked variables σ^* and y^* represent posterior standard deviation and mean value. Straight forward algebra and coefficients matching leads to general expressions for the posterior variance and mean values:

$$y^* = \mu + ((y - \mu)\sigma^2)/(1 + \sigma^2) \tag{2.17}$$

$$\sigma^{*2} = \sigma^2 - 1/(1 + \sigma^2), \tag{2.18}$$

with μ and σ being the mean value and standard deviation of the prior distribution. Note, that y^* gets larger or smaller compared to μ depending on the sign of $(y - \mu)$. The posterior value σ^* in 2.18 can only get smaller compared to the prior σ.

2.2 Bayesian Updating of an Application Case Using Benchmarks

The application of Bayesian statistics in the field of nuclear criticality safety is not new and for example used in SCALE's TSUNAMI [17]. Applying the above given formulas to criticality safety assessments is described in [7, 18, 16]. Assuming an application case with normal distributed $(N(\mu, \sigma))$ $p(k_{\text{eff}}^{\text{App}}) = N(k_{\text{eff}}^{App}, (\text{cov})^0)$ the prior (or knowledge) can be updated with suitable experimental data or benchmarks. The measurements can

be expressed as likelihood functions $p(v|k_{\text{eff}}^{\text{App}}) \propto \exp\left[(Uk_{\text{eff}}^{\text{App}} - v)^T \text{cov}_v^{-1} (Uk_{\text{eff}}^{\text{App}} - v)\right]$, with the vector of measured values $v = \vec{v}$, U a rectangular $m \times n$ matrix and cov_v the covariance matrix of the measurements.

If we assume an application case A and some benchmark experiments B and their k_{eff} values k_{eff}^A and $k_{\text{eff}}^B = k_{\text{eff}}^{\vec{B}}$ as functions of the nuclear data α and system describing parameters $\vec{x}$, $k_{\text{eff}}(\alpha, \vec{x})$, one can define the following prior distributions:

$$K^0 = (k_{\text{eff}}^{0,A}, k_{\text{eff}}^{0,B}) \tag{2.19}$$

for the mean values of the k_{eff} prior distribution and

$$\text{cov}^0 = \begin{pmatrix} \text{cov}^{0,A} & \text{cov}^{0,AB} \\ \left(\text{cov}^{0,AB}\right)^T & \text{cov}^{0,B} \end{pmatrix} \tag{2.20}$$

for the prior covariance matrix. The covariance of the application case is denoted by $\text{cov}^{0,A}$, for the benchmark experiments by $\text{cov}^{0,B}$. The matrix or vector (depending on the number of application cases) $\text{cov}^{0,AB}$ is the covariance of the application case and benchmark experiments due to nuclear data. It represents the comparability of the neutron spectrum and can be viewed as a weighting function. Normalizing these values with the standard deviations would lead to correlation coefficients as given for example by SCALE's TSUNAMI c_k values.

The covariances due to system parameters are expressed as $\text{cov}^{\text{sys}} = \text{cov}^{\text{sys},B}$, assuming statistical independence between the system parameters of the application case and benchmark experiments.

Using some algebra and equations 1.17 and 1.18, we derive the same expressions as [7] for the posterior model parameters of the application case k_{eff} distribution function:

$$k_{\text{eff}}^{A,*} = k_{\text{eff}}^{0,A} + \text{cov}^{0,AB} \left(\text{cov}^{0,B} + \text{cov}^{\text{sys},B}\right)^{-1} \left(k_{\text{eff}}^{B,\text{exp}} - k_{\text{eff}}^{B,\text{th}}\right) \tag{2.21}$$

$$\text{cov}^{A,*} = \text{cov}^{0,A} - \text{cov}^{0,AB} \left(\text{cov}^{0,B} + \text{cov}^{\text{sys},B}\right)^{-1} \left(\text{cov}^{0,AB}\right)^T \tag{2.22}$$

The updated knowledge of the application case mean value in 2.21 depends on the prior $k_{\text{eff}}^{0,A}$, the similarity of the application case and benchmark experiments $\text{cov}^{0,AB}$, the covariance of the used benchmark experiments due to nuclear data uncertainties $\text{cov}^{0,B}$ and due to uncertainties of the model parameters $\text{cov}^{\text{sys},B}$ and the difference between

the experimental and theoretical k_{eff} values, $k_{\mathrm{eff}}^{B,\exp} - k_{\mathrm{eff}}^{B,\mathrm{th}}$. The application case posterior mean $k_{\mathrm{eff}}^{A,*}$ can become smaller or larger in comparison to the prior depending on the sign of the difference between experimental and theoretical benchmark k_{eff} values.

3   Generation of Covariance Matrices

In the preceding section 2.1 we presented the posterior distribution characteristics and their dependence on covariance data. In the following section we describe in detail the generation of integral experiment covariance matrices $\mathrm{cov}_{\mathrm{sys,B}}$ based on publicly available data sets of critical experiments. The generation of covariance matrices due to nuclear data uncertainties will be presented, but not discussed in greater details, since it is already described in the literature, e.g. in [19, 20, 21].

The primary source for publicly available experimental data of critical experiments is the *International Handbook of Evaluated Criticality Safety Benchmark Experiments* (ICSBEP) [1]. It contains criticality safety benchmark specifications that have been derived from experiments performed at various critical facilities around the world. The evaluated criticality safety benchmark data are given in nine volumes and contain 567 evaluations with benchmark specifications for 4 874 critical, near-critical or sub-critical configurations.

The evaluated data given in the ICSBEP is widely used for validation purposes. Subsequently we will consider evaluated data for arrangements of low enriched uranium fuel rods and plutonium solution, both for thermal neutron spectra. The experiments were conducted in series, meaning that they were repeated with a slight modification of the setup. This could be for example the alteration of empty spots in a grid of fuel rods or a different pitch between the fuel rods. This procedure introduces correlation effects in the resulting k_{eff} values. Based on the data we will generate the covariance data using Monte-Carlo sampling techniques.

3.1   Monte-Carlo Sampling

Covariance data can be derived either by Monte-Carlo methods or based on linear perturbation of model parameters. The covariance matrix Σ of a parameter, e.g. $k(x)$ with $x = \vec{x}$ can be defined as

$$\Sigma_k = S\Sigma_x S^T \tag{3.23}$$

with S being the sensitivity matrix of the k_{eff} values with respect to system parameters x at $x = x_0$, and Σ_x the covariance matrix of x. If the system parameters $x1, \cdots, x_n$ are

chosen to be independent, Σ_x takes diagonal form:

$$\Sigma_x = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_n^2). \tag{3.24}$$

Following the Monte-Carlo sampling approach for the system parameter x and the corresponding probability density functions, a Monte-Carlo transport code, e.g. KENO-V.a can be used to calculate multiple neutron multiplication factors k_i. After a sufficient number of Monte-Carlo cycles, the covariance matrix Σ_k can be approximated by the sample covariance matrix

$$\hat{\Sigma}_k = \frac{1}{n_{MC}-1} \sum_{i_{MC}=1}^{n_{MC}} (k_{i_{MC}} - n_{MC}^{-1} \sum_{i_{MC}=1}^{n_{MC}} k_{i_{MC}})(k_{i_{MC}} - n_{MC}^{-1} \sum_{i_{MC}=1}^{n_{MC}} k_{i_{MC}})^T$$
$$\tag{3.25}$$

Using the Monte-Carlo approach can have several advantages. It is easy to adapt, since no sophisticated mathematical tools or knowledge is needed. Since it does not rely on further simplification e.g. taking only 1st order perturbations into account, non-linear dependencies of a varied parameter and the result are included. A disadvantage of Monte-Carlo sampling Methods compared to linear perturbation approaches can be the number of calculations needed to get a statistically satisfying result. The summed calculation time for the repeatedly evaluated calculation models with varying starting conditions can reach large values, depending on the model and number of varied input parameters.

3.2   SUnCISTT

In this work, we applied SUnCISTT [22] to execute and evaluate the Monte-Carlo sampled SCALE 6.1.2 criticality calculations. The GRS development SUnCISTT (Sensitivities and Uncertainties in Criticality Inventory and Source Term Tool) is a modular, easily extensible abstract interface program, designed to perform uncertainty and sensitivity analysis in the field of criticality safety. It couples different criticality and depletion codes commonly used in nuclear criticality safety assessments to the well-established GRS tool SUSA [23] and various Python packages (e.g. NumPy, SciPy, Matplotlib) for sensitivity and uncertainty analysis. SUnCISTT handles the complex bookkeeping that arises in the transfer of the generated samples into valid models of a given problem for a specific code. It generates and steers the calculation of the sample input files for the used codes. The computed results are collected, evaluated, and prepared for the statistical analysis. A detailed description of SUnCISTT is given in [22] and references therein. Here we just review the two modes *prepareSamples* and *CollectResults*, depicted in fig. 3.5 and 3.6 respectively.

To perform a Monte-Carlo sampling analysis of a given mathematical model, SUnCISTT needs information about the generated samples, the computational model to be analyzed and the specific input file requirements of the code to be executed. For the generated samples, an ASCII formatted list of the statistical varied input parameters has to be provided. This list can be generated by any suitable program. In fig. 3.5 the GRS program SUSA is depicted. For evaluations of large data sets, e.g. several experiments with a large number of samples and input variables we rely on a statistic program based on python. To generate the random numbers used for the Monte-Carlo sampling we used a Mersenne Twister pseudo-random number generator of the NumPy class *numpy.random.RandomState* [24]. Within this class a compatibility guarantee is given, meaning that a fixed seed initializing the random number generator and a fixed series of calls to 'RandomState' methods using the same parameters will always produce the same results up to round-off errors. Practically this means, that one does not need to store the entire ASCII list of random numbers, but for each parameter the distribution and its characteristics and the corresponding random seed. From the SCALE input file of the nominal case, a template file is derived in which user defined keywords replace the nominal values of the uncertain parameters. The third file to be provided in the SUnCISTT mode *prepareSamples* is a configuration file that sets the information of the other files into relation. With the given information, the desired number of individual input files with the statistical varied input parameters defined in the sample list are generated and the execution of them is steered. SUnCISTT also collects the calculated individual results and prepares them for further analysis. By default, SUnCISTT produces an ASCII formatted result file that can be transferred to SUSA or Python based scripts for the statistical evaluation. For further visualizations and analysis optional files for ROOT [25] are generated.

3.3   Example 1: Modeling Experimental Data

The questions arising in the field of determination and handling of integral experimental covariance matrices in the process of code validation for criticality safety calculations are discussed in the Expert Group on Uncertainty Analysis for Criticality Safety Assessment (UACSA), a sub-group of the Working Party on Nuclear Criticality Safety (WPNCS) of the Nuclear Energy Agency (NEA) within the Organization for Economic Co-operation and Development (OECD). Some of the actual questions which arose recently are: How to treat given sets of similar experimental data without knowing all exact statistical depen-
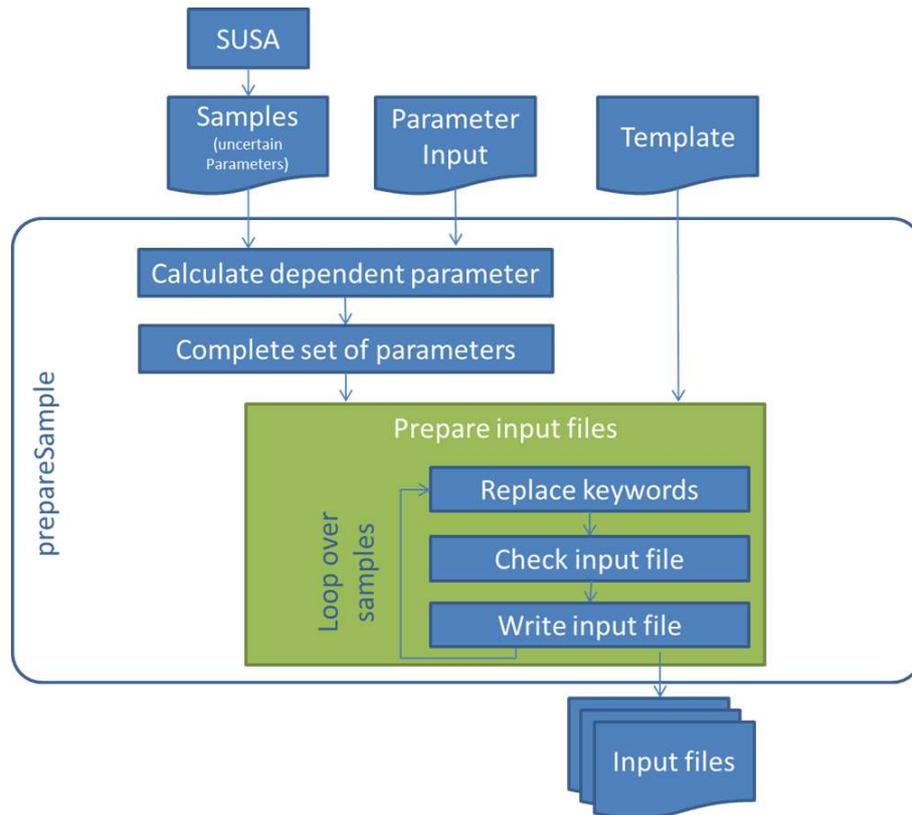
Fig. 3.5   Sketch of SUnCISTT sequences to generate samples. The mode *prepa-reSamples* needs a list of samples for the statistical independent input parameters, the parameter definitions, and a template file of the code to be executed. It then completes the set of parameters by deriving the depending parameters. By replacing the keywords in the template file with the parameters, the desired number of input files are generated.

dencies; and further, what are the implications on modeling these experiments in a code validation procedure regarding the consideration of the complete integral experimental correlation or covariance matrices?

In the following we address these questions by following parts of the group's proposal for a benchmark called *Role of Integral Experiment Covariance Data for Criticality Safety Validation* [11]. In contrast to the benchmark proposal we focus on a reduced number of experiments but a total of nine different modeling approaches. With the following analysis we show the effect of different modeling approaches for the same set of experimental data on the resulting integral covariance or correlation matrices.

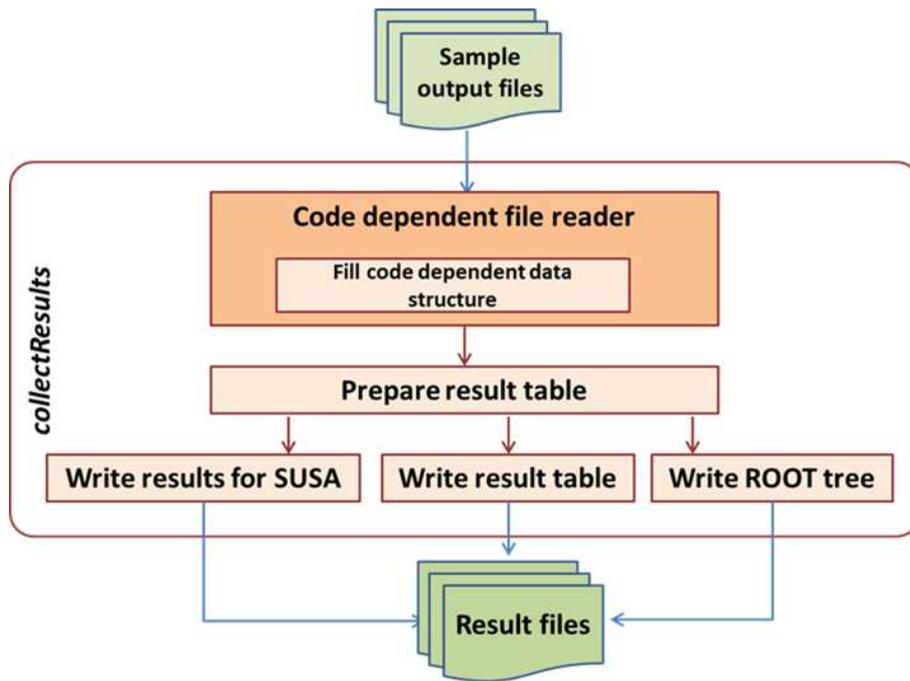As already mentioned, correlated data can arise if different experiments share parts of

Fig. 3.6   Sketch of SUnCISTT sequences to collect individual sample results and create integral result files. The mode *collectResults* needs some information about the code dependent structure of the generated output files to create results tables and ROOT-trees.

the experimental setup, measurement systems, or other relevant parameters. Some experiments described in the ICSBEP are not performed as single experiments, but slight variations of a setup were repeatedly investigated and published as a series of the same experiment. This is e.g. the case for LEU-COMP-THERM- 039 (LCT-39), where the number and locations of empty positions in a fuel rod grid were varied. In the following work we focus on the experimental data from experiments numbers 6, 7, and 8 from this series described in detail in [1, 11, 5] and references therein. The critical experiments consist of water moderated low enriched uranium fuel rods with a thermal neutron spectrum. The experimental setups are 22×22 arrays consisting of 363 (459, 448) fuel rods for experiment 6 (7, 8) and 121 (25, 36) empty positions, respectively. For further details we refer to [1, 5]. Clearly these experiments share certain components, and treating them as individual statistical independent data sets in the process of validation probably would not be appropriate. Hence, the determination of the integral covariance or correlation matrix of the experiments is a crucial step on the way to determine a bias of the calculated application case k_{eff}.

Following a Monte-Carlo Sampling approach, each value describing the experiment has

to be interpreted as a distribution function. This means in turn, that the definition and interpretation of the experimental parameters and their uncertainties is essential. It strongly depends on the quality of the experimental data and availability of precise uncertainty specifications. To circumvent the problem of determining suitable distribution functions for each parameter, we apply the ones proposed in the benchmark description, listed in tab. 3.1. All experimental parameters are supposed to follow either a uniform U(a,b) or normal distribution N(μ,σ). Assuming the three experiments LCT-39 6, 7, and 8 to be statistical independent gives a correlation coefficient close to zero. Results for this assumption are shown for the correlation of k_{eff} values calculated by KENO-V.a using the parameters given in tab. 3.2 for 250 Monte-Carlo samples for each experiment. The underlying model assumptions for the results of Fig. 3.7 are very simple and straight forward: It is assumed, that the fuel rods are all identical in composition and position within its unit cell. In consequence, the modeling of one experiment consists basically of a 22×22 array of identical unit cells for the fuel rods and the empty positions respectively.
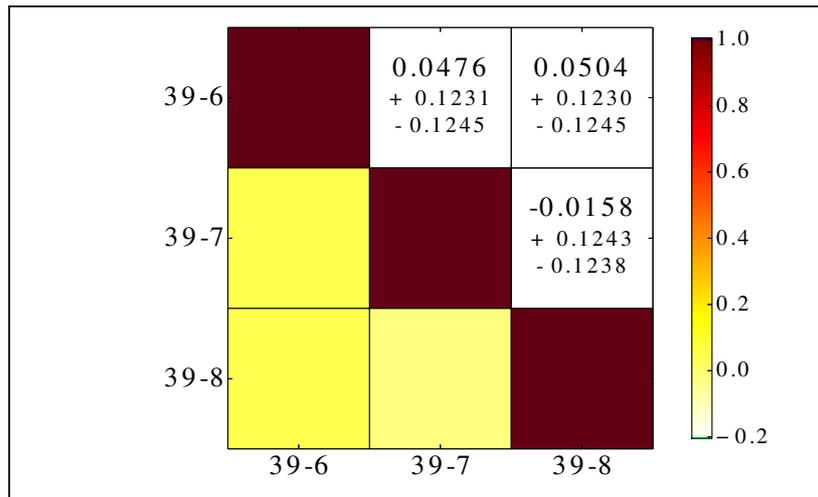

Fig. 3.7   Correlation coefficients of the experiments LCT-39 6, 7, and 8 assuming all fuel rods to be identical and no statistical dependence between each experimental setup. Shown are the cor-values and the upper and lower 95% CI.

3.3.1   Dependence on Modeling Choices

Having determined all relevant parameters and their distribution functions, a calculation model is built to calculate the neutron transport equations and determine the neutron multiplication factor. Obviously the model should be as close as possible to the experimental

Tab. 3.1   All model parameters and their distribution characteristics, following the suggestions of the benchmark proposal [11].

Model parameters	Type of variation	Distribution functions
Fuel diameter [cm]	depends on scenario	N(0.7892, 0.0017)
Fuel lengths [cm]	depends on scenario	N(89.7, 0.3)
Fuel density [g/cm^3]	depends on scenario	N(10.38, 0.0133)
Fuel content 234U [At.-%]	depends on scenario	N(0.0307, 0.0005)
Fuel content 235U [At.-%]	depends on scenario	N(4.79525, 0.002)
Fuel content 236U [At.-%]	depends on scenario	N(0.1373, 0.0005)
Boron concentration [atom/barn×cm×10^{-8}]	depends on scenario	N(6.9037,0.8)
Critical water height [cm]	individual	N(μ, σ) dep. on experiment
Angle of fuel rod	individual	U(0, 2π)
Offset of grid hole x [cm]	individual	N(0, 0.00742)
Offset of grid hole y [cm]	individual	N(0, 0.00742)
Hole diameter [cm]	depends on scenario	N(0.0105, 0.0085)
Inner cladding diameter [cm]	depends on scenario	U (0.81, 0.83)
Cladding thickness [cm]	depends on scenario	U (0.055, 0.065)

setup to get reasonable results. However, in the statistical interpretation of experimental series, the available data might leave some freedom of choice. The results shown in fig. 3.7 represent a model simplification of the experimental setup by assuming all fuel rods in one sample to be identical. However, one can argue, that due to manufacturing tolerances of the experimental equipment individual fuel rods may vary in both, their individual composition and position within the unit cell. The position of the fuel rod in the unit cell is then limited by the grid hole. For some simplicity we assume the fuel rod to be always vertical, meaning a 90 degrees angle to the horizontal plane. The modeling approach for the fuel rod displacement is depicted in fig. 3.8. For the modeling of the experiments in KENO-V-a, this implies each fuel rod to be simulated within its own unit cell, which we assume to have fixed dimension for all fuel rods. According to fig. 3.8, position of the grid hole might be displaced from the center of the unit cell by δ_x and δ_y in x and y direction. The center of the fuel rod itself might again be displaced in x- and y-direction, denoted by the radial displacement R and angle θ. In our modeling approach R is indirectly defined by the assumption that the fuel rod is in contact with the grid hole.

Tab. 3.2 Used codes and cornerstones of the calculations. KENO-V.a is taken from the CSAS5 sequence of SCALE 6.1.2.

Code	Parameter	Value
KENO-V.a	Nuclear data library	ENDF/B-VII (ce)
	Neutrons per generation	10,000
	Skipped generations	500
	σ_{MC}	5×10^{-4}(Sc. A to D); 1×10^{-4}(Sc. E to H)
SUnCISTT	Number of samples	250
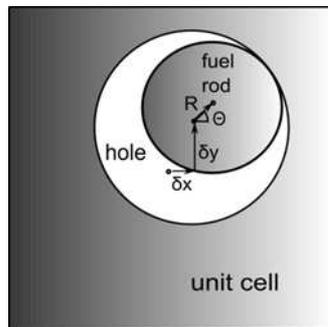
Fig. 3.8 Modeling approach for the displacement of the grid hole and fuel rod in a unit cell (not to scale). The center of the hole might vary from the center of the unit cell by δx and δy, respectively. The center of the fuel rod can vary by the distance R under the angle θ. As a boundary condition, the fuel rod has always contact with the grid.

We chose eight different modeling approaches, scenarios A to H, depending on assumptions on the fuel similarity and position of each single fuel rod. Scenarios A to E assume the fuel to be identical for all fuel rods in all experiments. One can argue that this might be a reasonable approximation, based on the assumption of a very accurate fuel fabrication process with only tiny tolerances. However, scenarios F and G assume a set of the maximal needed fuel rods (484 for the 22×22 grid array) being statistical independent. These fuel rods are placed for all experiments in a fixed position for scenario F or randomly for each experiment in scenario G, see fig. 3.9. Finally, scenario H assumes all fuel rods in every experiment to be statistical independent. The statistical dependence of the fuel between two experiments decreases from scenario E to H. The results shown in fig. 3.7 assume the same modeling assumptions as scenario A but any correlations were neglected, meaning no statistical dependence between each experimental setup. This scenario is named 'NoCor'.

Tab. 3.3 Modeling assumptions for fuel rod geometries and compositions of one sample for each different scenario. The variation of fuel in the last column means the variation of the diameter, length, density and enrichment of the fuel as well as the boron impurity. Scenario NoCor (scenario H) is identical to A (scenario G), except for neglecting statistical dependencies between experiments.

Scenario	Grid hole displ.	Grid hole diameter	Inner clad diam	Cladding thickness	Fuel variation
NoCor	centered	shared	shared	shared	shared
A	centered	shared	shared	shared	shared
B	δ_x, δ_y	shared	shared	shared	shared
C	δ_x, δ_y	individual	shared	shared	shared
D	δ_x, δ_y	individual	individual	shared	shared
E	δ_x, δ_y	individual	individual	individual	shared
F	δ_x, δ_y	individual	individual	individual	484 fixed pos.
G	δ_x, δ_y	individual	individual	individual	484 random pos.
H	δ_x, δ_y	individual	individual	individual	individual

The analysis allover required a total of 6,750 SCALE inputs with up to 20,000 lines per input file. The calculations were performed using a total of 60,000 CPU-h and 882 TByte-h. The results then were processed and statistically analyzed using SUnCISTT.

The resulting k_{eff} values for each experimental data set and modeling scenario are shown in fig. 3.10. We found a good agreement within the 2-σ range of the experimental data $k_{\mathrm{eff}}^{\mathrm{exp}} = 1.0$ (± 0.0012 for exp. 7,8) (± 0.0009 for exp. 6) given in [1]) and our results. The SCALE calculations with the applied continuous energy library ce_v7_endf (based on ENDF/B-VII) in the CSAS5 sequence systematically underestimates k_{eff}, which is a known effect for low enriched uranium setups [26]. Scenarios E to H in which the individual variations of the parameters partly cancel out each other, have significant lower error bars. The larger error bars of the Monte-Carlo approach of Scenario A in comparison to the error propagation approach done in the ICSBEP Handbook are not attributed to a general difference between the two methods. They rather arise from a different interpre-
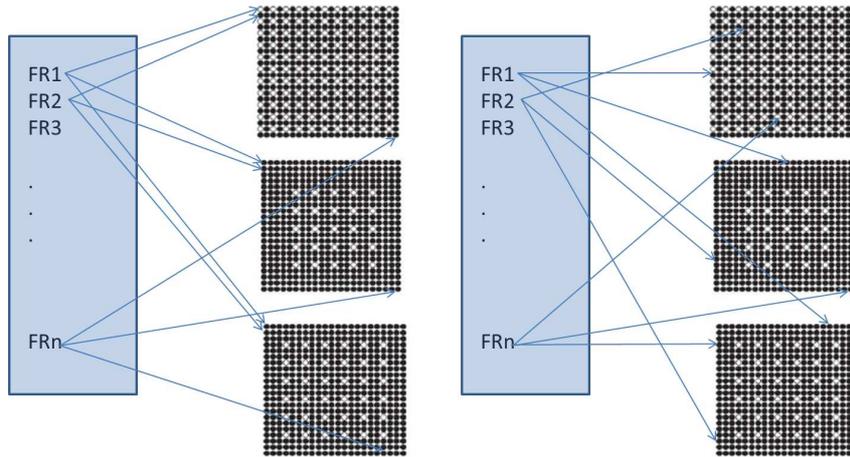
Fig. 3.9 Sketch of modeling approaches for scenarios F (left) , G and H (right). The blue boxes represent the fixed library of individual generated fuel rods. The black and white squares represent the three experimental setups LCT-39 6, 7, and 8 from top to bottom. A black dot represents a fuel rod, a white one an empty spot. The simulation corresponding to the left part of the picture assumes a fixed position for each fuel rod in each experimental setup. E.g. FR1 is always in the top left spot for every experiment. The right part of the figure depicts the assumption of each fuel rod being randomly placed in the grid for each experiment.

tation of the system parameter uncertainties. In the original experiment description [27] the uncertainty of the inner cladding diameter ($\pm$ 0.01 cm) and the cladding thickness ($\pm$ 0.005 cm) are reported to be independent. The uncertainty of the outer cladding diameter is obtained by error propagation. In the ICSBEP evaluation, the uncertainty of the cladding thickness is split equally between inner and outer diameter. This results in an uncertainty of the outer cladding diameter of $\pm$ 0.0025 cm, which reduces its impact on the uncertainty of k_{eff} significantly. The original evaluation assumes further a Gaussian distribution by dividing the half tolerance by $\sqrt{3}$. The resulting distributions for the outer cladding diameter and their impact on k_{eff} are different for both considerations, as shown in figures 3.11 and 3.12. Fig. 3.11 shows the distribution of the outer cladding diameter. Blue represents the distribution with parameters from the original literature, red with parameters given in the handbook. Fig. 3.12 shows the resulting distributions of k_{eff}. The black curve is the experimental target value of k_{eff} = 1.000 with a standard deviation (σ = 0.0012) deduced in [1] by linear error propagation from the experimental uncertainties. The blue curve shows the calculated distribution from a Monte-Carlo approach with pa-
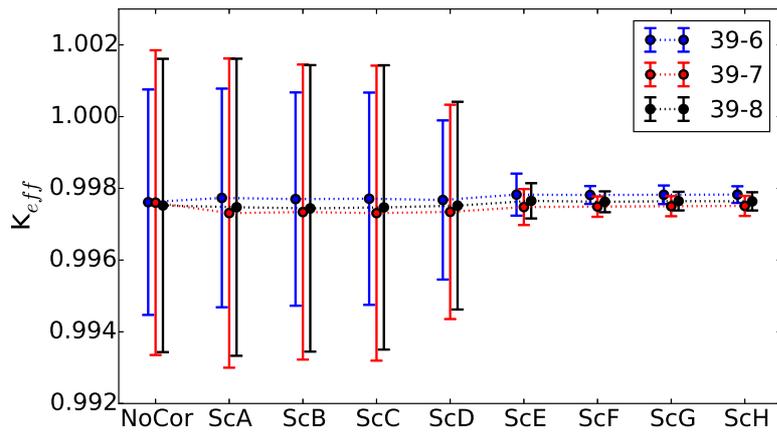
Fig. 3.10 Resulting k_{eff} values for 250 Monte-Carlo samples for each experiment and scenario. The error bars indicate the standard deviation. While the nominal values remain fairly constant for all scenarios, the standard deviation decreases significantly from scenario D to F.

rameters from the original literature. The red curve shows the result with the information from [1]. The resulting uncertainty agrees reasonable well with the one from error propagation in the handbook. However, the values based on the original literature were used for the following analysis.

In fig. 3.13 nine colored plots are shown for the cor values of the experiments LCT-39 6, 7, and 8 as well as the cor value and the 95% confidence interval. The results show correlation coefficients around 0 for the scenarios NoCor and H, as expected, since there are no relevant parameters with shared values between the individual experiments. Note that the difference between the NoCor and H scenario is the variation of the fuel rods: In contrast to NoCor, in scenario H each fuel rod in each experiment is simulated individually and statistically independent. This difference is mapped in the sensitivity plots in fig. 3.15 and 3.16 which show the correlation coefficients of each parameter with the resulting k_{eff}. Note, that the performed sensitivity analysis shows the impact of the actual variation of each parameter on the k_{eff} uncertainty. However, we do not perform a sensitivity analysis by varying only one parameter at a time. This means, our sensitivities depend on the chosen distribution functions and their characteristics. Changing these assumptions in our approach might lead to a different sensitivity profile. This approach was chosen since we are interested in determining the contribution of each varied input parameter on the uncertainty of k_{eff} for given modeling assumptions.

23
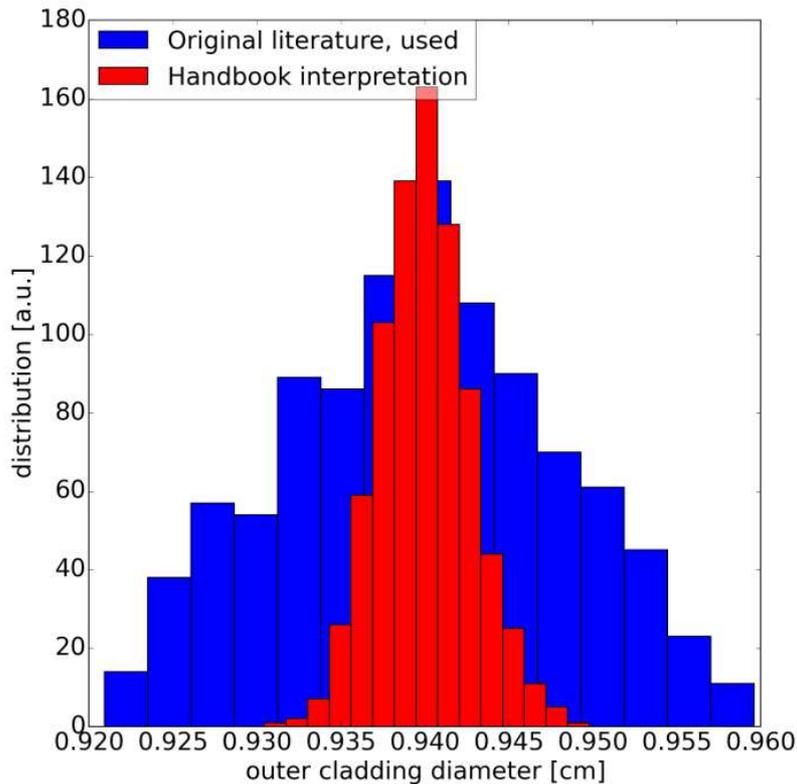
Fig. 3.11 Comparison of the distribution functions for samples of the outer cladding diameter. The red distribution follows the interpretation of [1], the blue distribution follwos the original published experimental data. The distribution functions are both normal distributed with the same mean value, but different σ's.

While for the scenario NoCor the most relevant parameters are the cladding inner radius and thickness, the only important parameter for scenario H is the critical water height. It is notable, that in this case the different interpretations of the given experimental data lead to comparable cor values but totally different sensitivity profiles. The highest correlation coefficients for scenario NoCor are the ones for the cladding inner diameter and thickness, and for the radius of the fuel. The only dominant parameter for scenario H is the critical water height. Scenarios A, B, C and D show all cor values close to 1 with only little deviations between the different cor values of the scenarios. The cor values for scenario A to C are even the same within the 95% confidence interval. Their corresponding sensitivity profiles show huge similarities: The three largest cor values are the cladding inner radius and thickness and the radius of the fuel. For scenario B and C the U-235 weight-% plays a more prominent role. The sensitivity profiles for scenario D show a different behavior since the inner cladding diameter here is varied individually for each fuel rod. The leading contribution to the sensitivity profile now solely results from the cladding
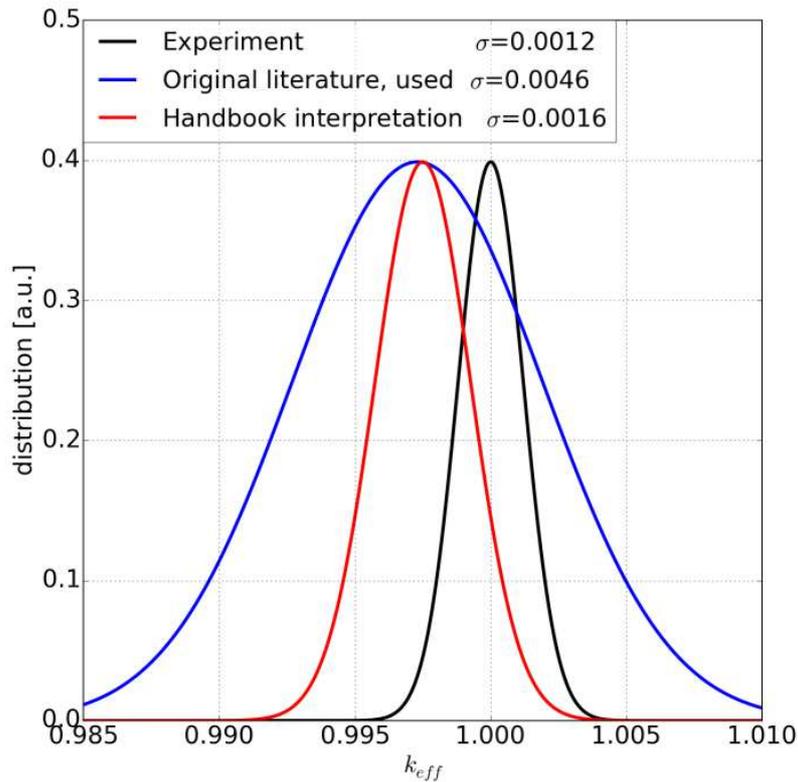
Fig. 3.12 Impact of the two different distribution functions for the cladding outer diameter on the calculated k_{eff} distribution compared to the experimental values (black). The distribution function based on the original values (blue) result in a larger uncertainty compared to the values from [1] (red).

thickness.

Scenario E shows correlation coefficients of approximately 0.75 between the experiments 6 and 7 and 0.7 between experiments 7 and 8. The difference between the cor values is due to lower number of fuel rods in experiment 6 compared to 7 and 8. Thus, the individual variation of the cladding inner radius and thickness for each fuel rod affects the correlation coefficient of the experiments 7 and 8 more. The corresponding sensitivity profile shows the fuel radius as the leading parameter. A mild impact is shown by the fuel density and critical water height (additional the weight-$\%$ for U-235 and the fuel height for LCT-39 7).

Scenarios F and G show significantly smaller correlation coefficients between the experiments. The difference to Scenario E is that now also the fuel content of each fuel rod is varied individually. This can be seen in the sensitivity profile of both scenarios in fig. 3.16, where the dominant parameter is the critical water height.
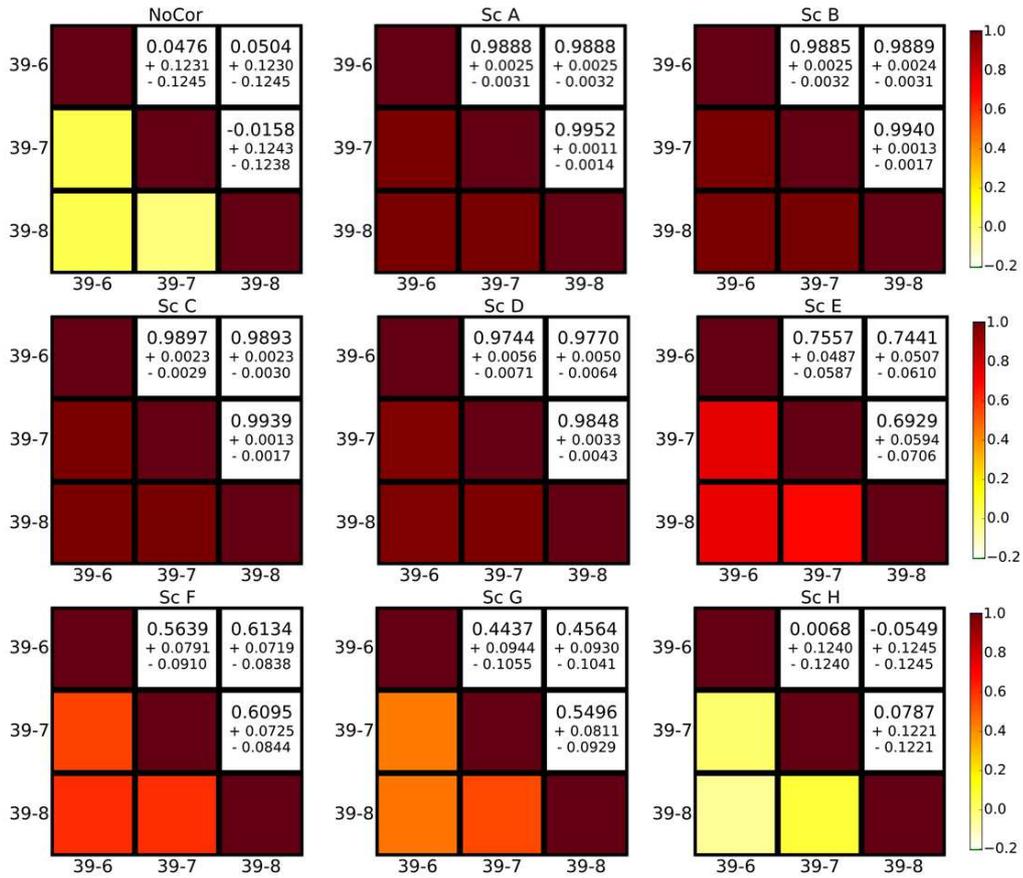
Fig. 3.13 Matrices of the correlation coefficients for the 9 scenarios. The upper part of each matrix shows the cor values and the ranges for the 95% confidence interval. The lower part shows the nominal cor values color coded. The scale varies from dark red for cor = 1 to white for cor = -0.2. The resulting correlation coefficients for scenario NoCor and H are 0 within the 95% confidence interval. Scenarios A to D show cor values close to 1, while scenarios E to G show a slight decreasing of cor from approximately 0.75 to 0.5.

3.3.2   Dependence on Neutron Transport Calculation

It is obvious, that the Monte-Carlo uncertainty $\sigma^{MC}_{k_{\text{eff}}}$ of k_{eff} has to be chosen to be significantly smaller compared to the experimental uncertainty. Otherwise, the effect of correlated data due to shared parameters in the experiment descriptions will remain hidden and the derived correlation coefficients will be too low. To illustrate this effect, we performed further SUnCISTT calculations using the CSAS5 sequence of SCALE 6.1.2 for the above described Scenarios F and G of the experiments LCT-39 6 to 8 (250 samples each), but varied $\sigma^{MC}_{k_{\text{eff}}}$. The latter was set to values of 1×10^{-3}, 5×10^{-4}, 1×10^{-4}, 5×10^{-5} and
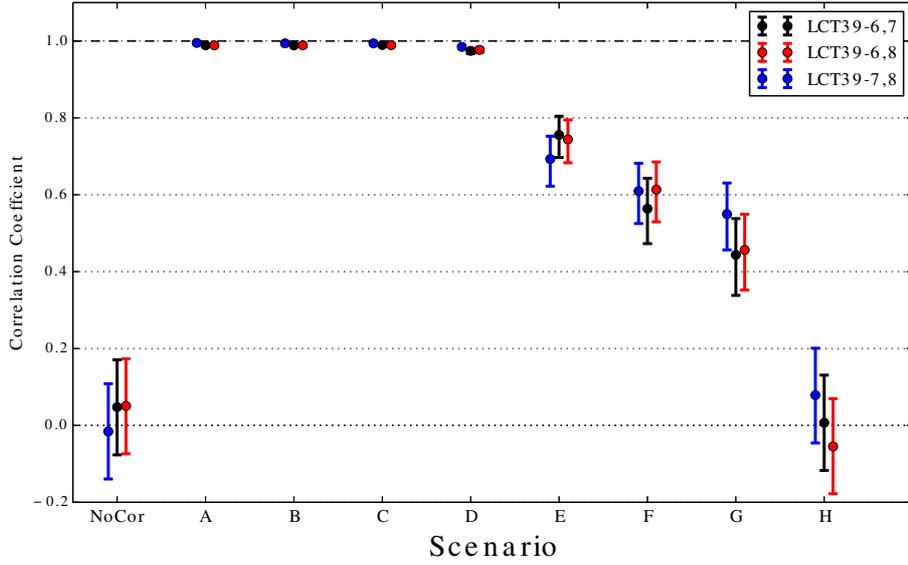
Fig. 3.14   Correlation coefficients for the pairs of experiment 6 and 7 (black), 6 and 8 (red), and 7 and 8 (blue) of Series LCT-39 for the 9 scenarios. The error bars represent the 95% confidence interval. Scenarios A, B, C, and D show high correlation coefficients, all close to 1. For scenarios E to G the coefficients decrease to approximately 0.7, 0.6, and 0.5 respectively. The correlation coefficients for scenario H are 0 within the 95% confidence interval, like the ones for the NoCor scenario.

additionally for Scenario F[1] to 2.5×10^{-5}. The results shown in fig. 3.17 and 3.18 illustrate the increasing of the correlation coefficient cor for the three experiments from around 0.1 to 0.2 for $\sigma_{k_{\mathrm{eff}}}^{MC} = 1 \times 10^{-3}$ to cor ≈ 0.7 (Scenario F) and cor ≈ 0.55 (Scenario G) for $\sigma_{k_{\mathrm{eff}}}^{MC} = 2.5 \times 10^{-5}$.

For values smaller than $\sigma_{k_{\mathrm{eff}}}^{MC} = 1 \times 10^{-4}$, the 95% confidence interval overlap and the change of mean values of the correlation coefficients is small compared to the change from $\sigma_{k_{\mathrm{eff}}}^{MC} = 5 \times 10^{-3}$ to $\sigma_{k_{\mathrm{eff}}}^{MC} = 1 \times 10^{-4}$.

---

[1] We did not perform the same calculation for Scenario G since it would have needed an estimated calculation of far more then 60k CPU-h and the expected result would only differ little from $\sigma_{k_{\mathrm{eff}}}^{MC} = 5 \times 10^{-5}$.

Tab. 3.4 Parameters and key words used in fig. 3.15 and 3.16.

Key word	Parameter	Key word	Parameter
rad_CladIn	cladding inner raduis	w%_U235	weight-% U-235
thick_Clad	cladding thickness	w%_U236	weight-% U-236
rad_Fuel	fuel radius	height_Water	water height
height_Fuel	fuel height	rad_Hole	hole radius
dens_Fuel	fuel density	delta_Hole_X	δ_x
dens_B10	B-10 density	delta_Hole_Y	δ_y
w%_U234	weight-% U-234	angle_Rod	θ

3.3.3 Dependence on Sample Size

Another technical obstacle in calculating the correlation coefficient using a Monte-Carlo Sampling approach is to use a significant number of samples to ensure convergence. The problem is widely known and discussed broadly in the literature with various recommendations (see e.g. [28] and references therein). A rule of thumb is, that the larger the number of samples is, the more the results tend to minimize the probability of errors, maximize the accuracy of population estimates, and increase the possibility to generalize the results. Some authors recommend a minimum number of samples of 50, others of 400 [29], for a review see e.g. [30].

A strong limiting factor for the number of samples used for our investigations in this work is the calculation time for each experiment. Sometimes values of 20,000 CPU-h were reached to calculate a covariance for two experiments and 250 samples each. In theses cases calculations based on 500 or even 1000 samples seems to become almost impossible.

To investigate the problem of how many samples are needed to get reliable results, we generated 625 samples of each experiment described above for Scenario A. We performed calculations using SUnCISTT with CSAS5 sequence of SCALE 6.1.2 and a $\sigma_{k_{\text{eff}}}^{MC} = 1 \times 10^{-5}$. Following the evolution of the correlation coefficient with increasing number of samples lead to the results shown in fig. 3.19 and 3.20. The 95% confidence interval becomes broader, the smaller the absolute value of the correlation coefficient becomes. Or in other words, the closer the correlation coefficient is to 1, the smaller is
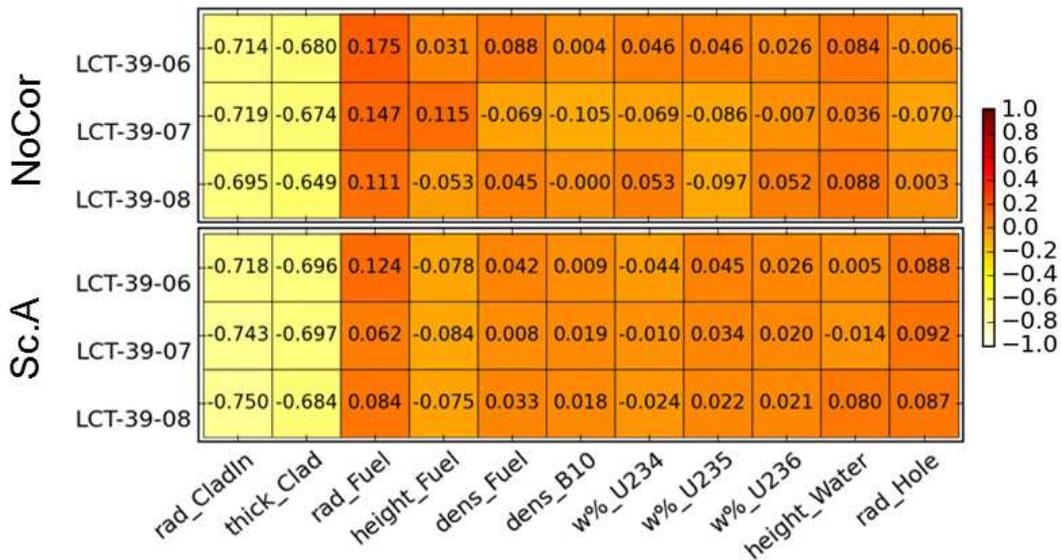
Fig. 3.15   Correlation coefficients of each individual model parameter listed in tab. 3.4 and the resulting 250 k_{eff} values for scenarios NoCor and A. The profiles are similar due to identical modeling assumptions. The only difference between the two scenarios is the assumption of correlations in scenario A. Note the different scale of the color coded representation of cor w.r.t. fig. 3.13.

the uncertainty due to sampling. This effect is due to the Fisher's z-distribution used to calculate the confidence intervals. It can already been seen, that the bigger the number of samples is, the less statistical noise corrupts the resulting correlation coefficient. For correlation coefficients close to the absolute value of 1, the results and converge faster and the confidence intervals remain almost constant. In fig. 3.20 it can be seen, that for the correlation coefficient of the experiments LCT-007 1 and 2 approximately 50 samples give already a very good and stable result, whereas the correlation coefficients for LCT-007 1 and 3, and 1 and 4 are still very unstable and have comparable huge confidence intervals.

If one wants determine a minimum number of samples needed, one has to investigate the evolution of the smaller correlation coefficients in more detail. In the following we will have a closer look on the evolution of the correlation coefficient of LCT-007 experiments 3 and 4 and its 95% confidence interval (fig. 3.21). The mean value remains fairly constant for sample numbers larger then 50. But the 95% confidence interval still allows positive correlation coefficients, whereas these values are excluded for lager values of samples numbers. Performing a fit of the upper and lower confidence levels each via exponential
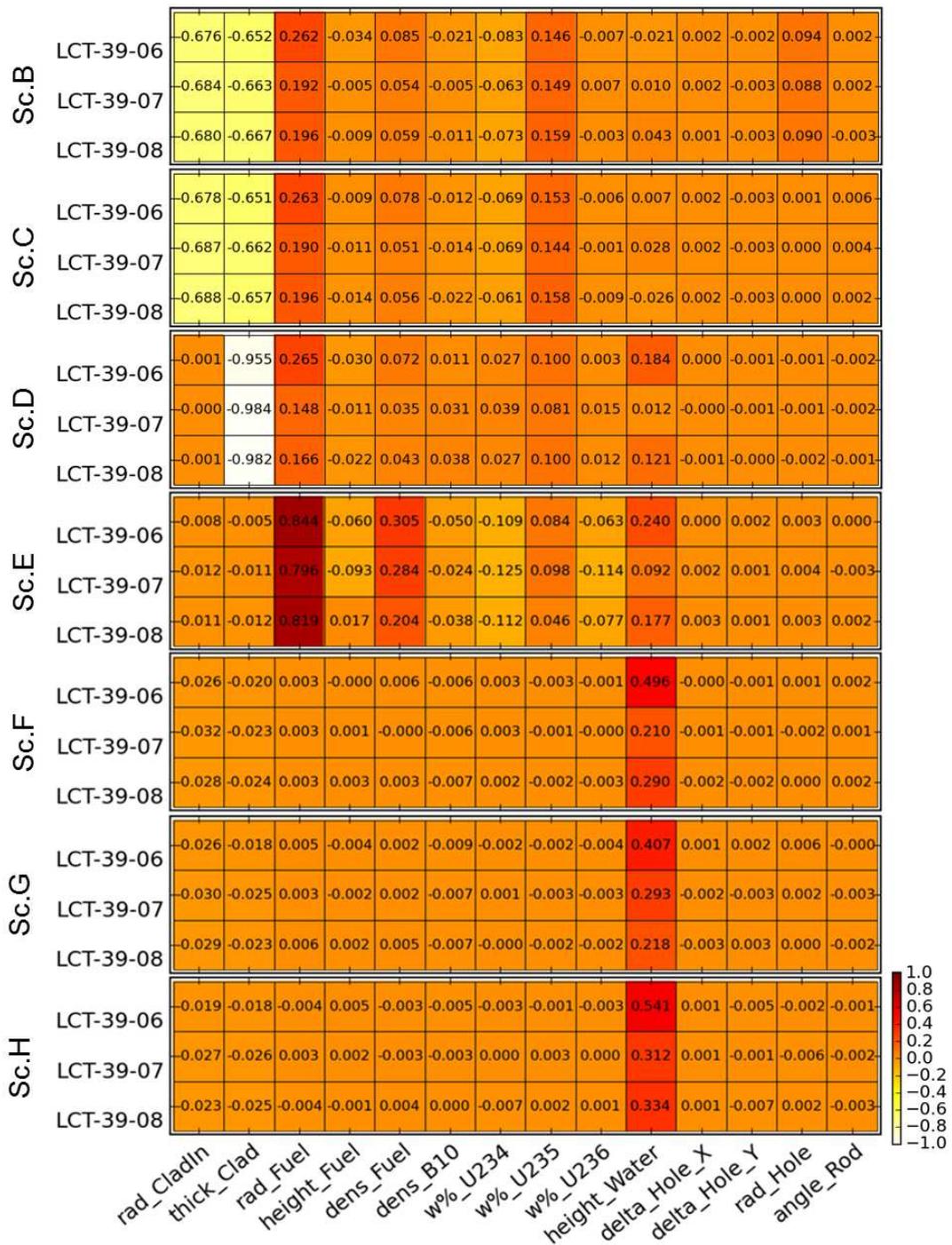
Fig. 3.16 Correlation coefficients for each model parameter and the resulting 250 k_{eff} values. From top to bottom the scenarios B to H are shown. The increased number of parameters compared to the scenarios NoCor and A is due to the variations of single fuel rods. The color coded representation of the cor shows values from -1 to 1.
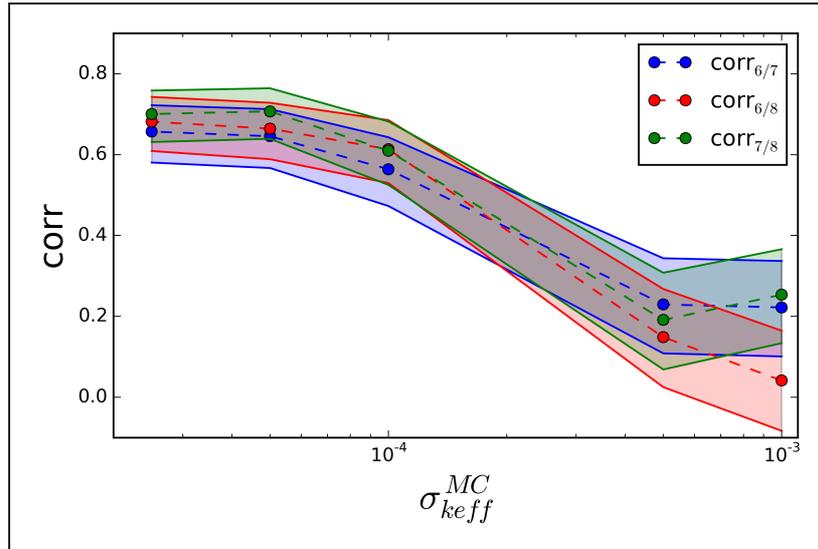
Fig. 3.17 Varying the Monte-Carlo uncertainty for the k_{eff} calculations in KENO-V.a for Scenario F and the resulting cor values (colored dots) and the 95% CI (colored area between solid lines). The c_k values increase from approximately 0 (correlations between experiments 6 and 8) and 0.2 (correlations between remaining experiments) for $\sigma_{k_{\text{eff}}}^{MC} = 1 \times 10^{-3}$ to approximately 0.7 for $\sigma_{k_{\text{eff}}}^{MC} = 2.5 \times 10^{-5}$.

fit functions $f(x) = a \times \exp[\text{bx}] + \text{c}$ one can identify the linear regime of the confidence interval in which it remains fairly constant. In our analysis shown in fig. 3.21 this regime starts at approximately 200 samples. Choosing 200 or more samples leads to acceptable results even for correlation coefficients around zero.

3.3.4   Summary and Discussion

We discussed different modeling approaches for a given set of experimental data, leading to different correlation coefficients and sensitivity profiles. Using a Monte-Carlo approach, we calculated 250 samples for each experimental setup and scenario to obtain the resulting k_{eff} values (fig. 3.10). We showed the impact of statistical parameters like sample size and the uncertainty due to the Monte-Carlo approach of the neutron transport code.

Within each scenario we calculated for each pair of experiments the corresponding correlation coefficient and the 95% confidence intervals of this coefficient (fig. 3.13, 3.14). We showed for each experiment and scenario the impact of the variation of each input parameter on the resulting k_{eff} by calculating the corresponding correlation coefficients
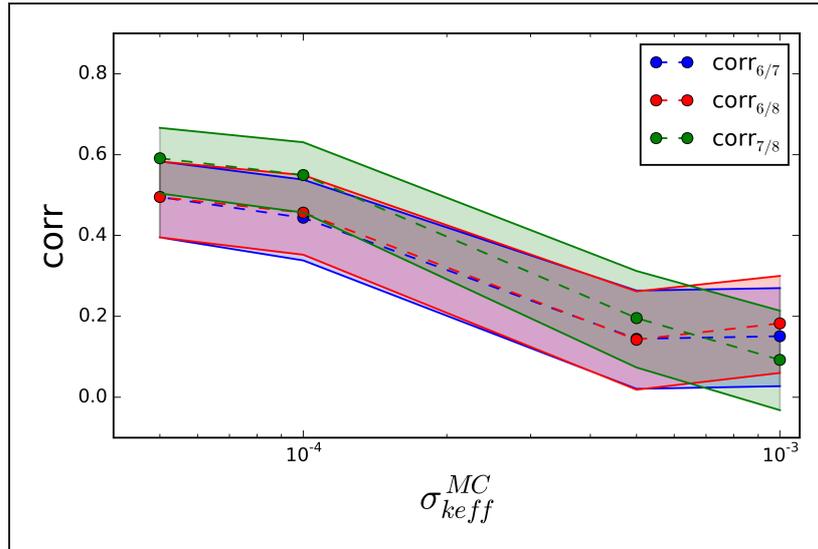
Fig. 3.18   Varying the Monte-Carlo uncertainty for the k_{eff} calculations in KENO-V.a for Scenario G and the resulting cor values (colored dots) and the 95% CI (colored area between solid lines). The c_k values increase from approximately 0.1 (correlations between experiments 7 and 8) and 0.2 (correlations between remaining experiments) for $\sigma_{k_{\text{eff}}}^{MC} = 1 \times 10^{-3}$ to approximately 0.6 and 0.5 respectively for $\sigma_{k_{\text{eff}}}^{MC} = 2.5 \times 10^{-5}$.

(fig. 3.15, 3.16). For the combination of water moderated, low enriched Uranium rods modeled with the criticality code KENO-V.a we found that the correlation coefficients between the k_{eff}'s of the experiments LCT-39 6, 7, and 8 varied between 0 and 1 within the 95% confidence interval. The modeling assumption leading to scenario A and NoCor are identical, except that NoCor neglects correlations completely. The same holds for scenarios G and H, which are identical, but H neglects the correlations between experiments due to fuel similarities.

Varying all geometrical parameters affecting the outer cladding radius for each fuel rod separately leads to a significant decrease of the resulting correlation coefficient compared to the results derived from the assumption of all fuel rods having identical geometrical parameters (however, scenarios B, C and D still comprise high correlations). We found a significant drop of the cor value from scenario D to E (3.14) as well as a significant drop of the k_{eff} uncertainty (fig. 3.10). The main contribution to the k_{eff} uncertainty in scenario D stems from the cladding thickness, which in scenario E plays no role since it is varied for each fuel rod individually (fig. 3.16).
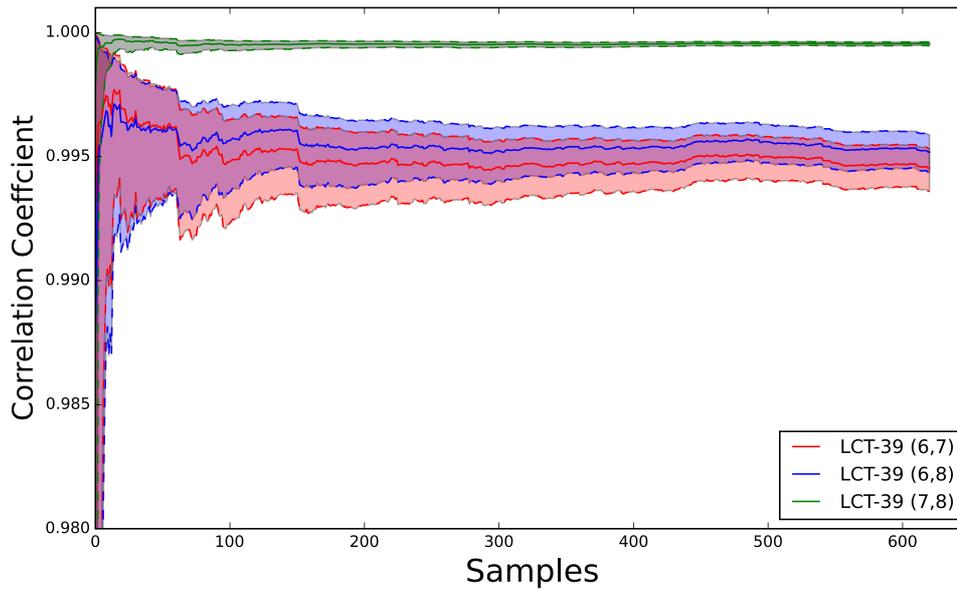
Fig. 3.19 Shown here are the evolution of correlation coefficients and the 95% CI for experiments LCT-39 6, 7, and 8 for increasing sample size. All cor values are close to 1, especially for experiments 7 and 8 (green), where the 95% CI remains fairly constant for sample size larger than 50.

Scenarios F and G show a further drop of the correlation coefficient, but within the 95% confidence interval the cor values of the two scenarios overlap (fig. 3.14). The difference of the two assumptions, knowing the exact position of each fuel rod for Scenario F or randomize their position in the grid for scenario G has a comparable smaller effect on cor than the assumption of a finite number of fuel rods. It is notable, that the sensitivity analysis shows the sole dependence of the k_{eff} uncertainty on the critical water height (fig. 3.16).

The different modeling assumptions might all be justified based on expert judgment. However, the sensitivity analysis reveals different sensitivity profiles, especially from scenario C to F. One could be tempted to choose the modeling assumptions based on the quality of the experimental data. As an example one could argue to choose scenario F or G, since the uncertainty of k_{eff} is much lower and the almost sole dependence of the k_{eff} uncertainty is on the critical water height. Following this argumentation, one could construct modeling assumptions based on the given experimental data to reduce uncertainties and to circumvent possible gaps in the data. But one has to be very careful with these options, and give very good arguments, why one chooses one scenario over ano-
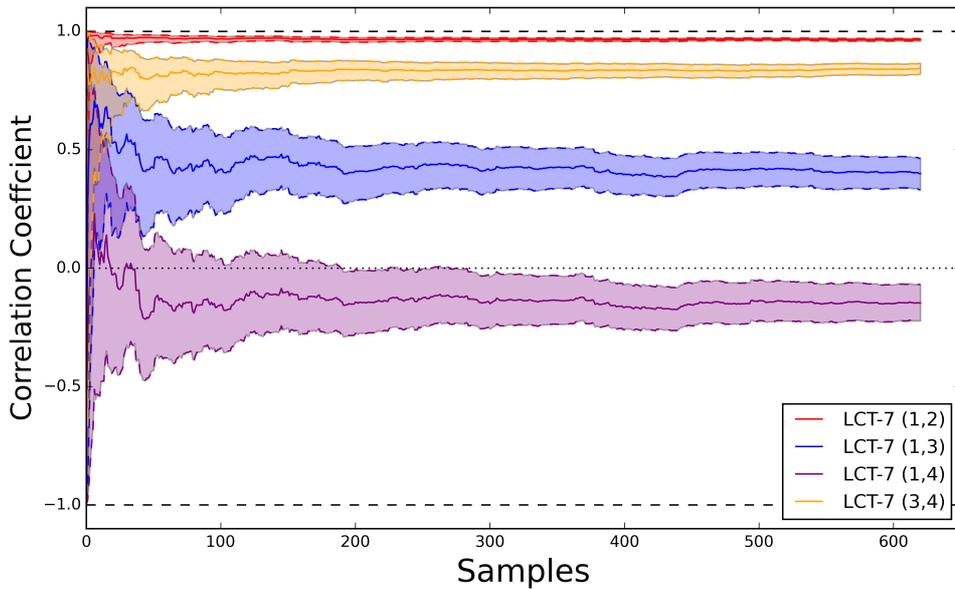
Fig. 3.20   Shown here are the correlation coefficients for pairs of experiments LCT-07 1, 2, 3, and 4. The CI band of the correlation coefficient of experiment 1 and 2 is small, since the values are close to 1. The CI bands for experiments 1 and 3 (blue), 1 and 4 (purple), and 3 and 4 (yellow) are wider, since their *cor* values are closer to zero. The number of samples needed for acceptable results w.r.t. the 95% CI increases for $|cor| \to 0$.

ther. The resulting covariance matrices directly influence the bias and its uncertainty, and thus the resulting upper sub-critical limit [7, 6, 8].

Using the covariance or correlation matrices for the purpose of validation or the determination of the upper sub-critical limit of an application case, the results can vary strongly, depending on the scenario. Following the argumentation of [6, 8], a rule of thumb is that the higher the correlation coefficient, the lesser information is available, and thus the upper sub-critical limit decreases. This means, that being not able to distinguish between the different scenarios and identify the correct one based on the available data, one would in this case take the results associated with the highest correlation coefficient to get a more conservative estimate of the bias in code validation or the upper sub-critical limit.

Note that the underlying data for the work presented is partly constructed and fictive as it is a part of a calculation benchmark exercise [11]. From the given data, any modeling assumption from scenario A to H could be justified. For further determination of the sce-
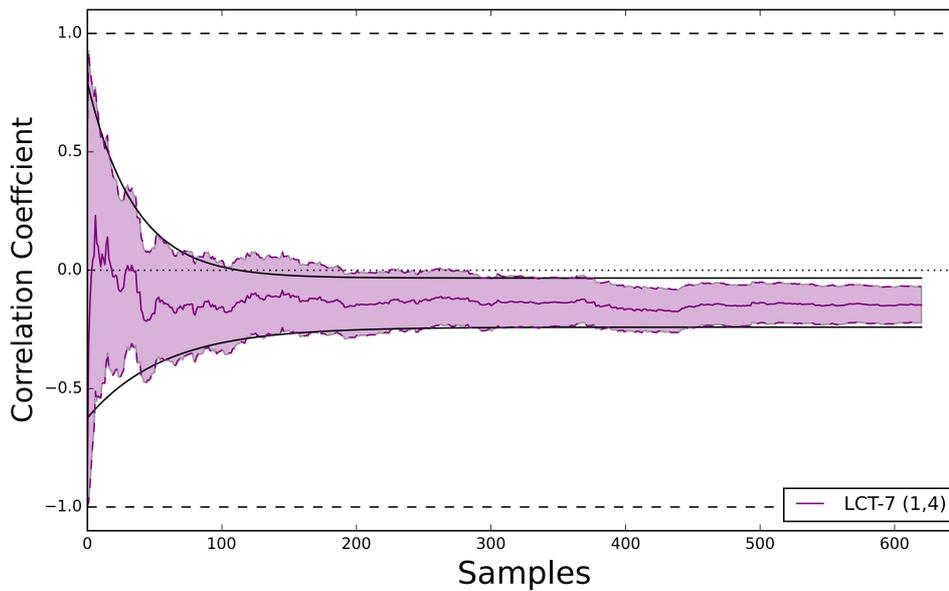
Fig. 3.21   Shown is the evolution of the correlation coefficient with increasing sample size for experiments LCT07 1 and 4 (solid purple line) and the 95% CI. The latter is fitted via two exponential functions for the upper and lower CI (black line). The fit is for sample sizes larger 200 almost linear, indicating an acceptable convergence of the CI.

narios one would need to know e.g. if the fuel content and geometric description for each fuel rod was identical or if it varied. The statements presented above thus are only valid for the combinations of code and experiments discussed here.

To derive more general statements, further investigations have to be carried out. On the other hand, it may be problem dependent if and to what extent the regard for correlations between benchmark experiments could influence the bias determination. A sufficient number of statistical independent data sets, e.g. for experiments conducted in different laboratories using different materials, can always circumvent the problem of the correct determination of integral experimental covariance data. However, the accurate consideration of correlated data seems to be inevitable if the experimental data in a validation procedure is limited. But even if one can avoid the determination of the accurate integral experimental covariance data due to statistically independent data sets, the selected modeling scenario should always be justified. The modeling assumptions have the potential to decrease the uncertainty of the resulting k_{eff} significantly.

3.4 Example 2: Deriving Statistical Data From Experiments

The previous subsections have shown, that for the determination of covariance and correlation coefficients, precise modeling of experimental data is essential. Another issue is related to the experimental data itself. Sometimes, not all relevant data needed for a Monte-Carlo Simulation is given in the experimental description. Without knowing all statistical dependencies and relevant parameters of the experimental data, they have to be guessed. Different assumptions, however, may lead to different covariance and correlation coefficients in criticality safety calculations. Most of the required information for calculations may be extracted from the documentation of experiments. However, depending on the level of detail, relevant information might be missing and has to be reconstructed or guessed by the evaluator, leaving some space for interpretation. Relevant parameters comprise experimental values (e.g. fuel radii, uranium concentrations and critical water heights), their uncertainties and types of uncertainty (as tolerance, measurement uncertainty or experience value), which translate to different uncertainty distributions. The most commonly used probability distributions of system parameters are normal or Gaussian and uniform distributions. The former arises e.g. from repeated measurements of a parameter or fitting procedures, while the latter applies for tolerances or in general if no specific information about the possible distribution is available. The experimental benchmark documentation usually provides all relevant experimental values, in many cases accompanied with the associated uncertainties. The types of uncertainties, however, are often missing and have to be chosen according to expert judgment.

In the following, the impact of different distribution functions of parameters describing an experiment is investigated by calculating the effective neutron multiplication factors k_{eff} and the correlation coefficients for selected benchmark experiments. Modeling a benchmark experiment, there is typically a mixture of normal and uniform distributed parameters. This case is in the following referred to as *expert judgment*. Additionally, two bounding cases are considered, where all parameters are either normal or uniform distributed. The database of benchmark experiments used in the following consists of LEU-COMP-THERM (LCT) experiments from the ICSBEP handbook [1]. All experiments are designed to be critical, i.e. the experimental effective neutron multiplication factors are $k_{\mathrm{exp}} = 1.0$.

A set of nine selected benchmark experiments from series LCT-06, 35 and 62 is used. All these experiments were carried out in the same laboratory sharing certain components.

In contrast to LCT-06 and 35, where only pellet fuel rods were used, a combination of powder and pellet fuel rods was used in LCT-62. Further details on the experiments can be found in the ICSBEP handbook.

Parameter	Value	Uncertainty	Distribution	Source	Comment
pellet fuel rods used in all three experiment runs					
U-234 (wt%)	0,008*U235	0,001*U235	$N(\mu,\sigma^2)$		assumed typicel fraction of (0,008±0,001) of U-235
U-235 (wt%)	2,600	0,02	$N(\mu,\sigma^2)$	LCT062, P. 28	uncert. is 1σ based on GE's "Quality Control Documents"
U-238 (wt%)	97,379				value is calculated
fuel density (g/cm^3)	10,368	0,04		LCT062, P. 27	calculated from measured wight, diameter and length of fuel stack
weight of UO2/rod (g)	1834,7	5,2	$N(\mu,\sigma^2)$	LCT062, P. 6	uncert. is 1σ of 198 fuel rods
fuel diameter (cm)	1,2502	0,0002	$N(\mu,\sigma^2)$	LCT062, P. 6	uncert. is 1σ based on measurements of new pellets in 1968
fuel stack length (cm)	144,15	0,3	$U(a,b)$	LCT062, P. 6	uncertanty is tolerance
clad outer diameter (cm)	1,4172	0,0011	$N(\mu,\sigma^2)$	LCT062, P. 6	uncert. is 1σ from measurements of 110 fuel rods
clad thickness (cm)	0,0762	0,0006	$N(\mu,\sigma^2)$	LCT062, P. 6	design thickness is 0,03 in
clad inner diameter (cm)	1,2648	0,002		calculated	calculated from measured outer diameter and clad thickness
bottom end plug (cm)	16,83				length
gap above fuel (cm)	2,54				length
top end plug (cm)	24,04				length
swaged fuel rods used in LEU-COMP-THERM-062 only					
U-234 (wt%)	0,008*U235	0,001*U235	$N(\mu,\sigma^2)$		assumed typical fraction of (0,008±0,001) of U-235
U-235 (wt%)	2,580	0,02	$N(\mu,\sigma^2)$	LCT062, P. 28	uncert. is 1σ based on GE's "Quality Control Documents"
U-238 (wt%)	97,399				calculated
fuel density (g/cm^3)	9,93	0,41		LCT062, P. 27	calculated from measured wight, diameter and length of fuel stack
weight of UO2/rod (g)	1829,2	17,3	$N(\mu,\sigma^2)$	LCT062, P. 6	uncert. is 1σ of 198 fuel rods
fuel diameter (cm)	1,270	0,025		LCT062, P. 6	uncert. is 1σ, calculated from outer diameter and thickness
fuel stack length (cm)	145,415	1,27	$U(a,b)$	LCT062, P. 6	uncertainty is tolerance
clad outer diameter (in)	0,556	0,0005	$U(a,b)$	LCT062, P. 6	uncertanty is tolerance, design value is 0,556 in
clad thickness (in)	0,028	0,005	$U(a,b)$	LCT062, P. 6	uncertanty is prob. tolerance, design thickness is 0,028 in
clad inner diameter (cm)	1,27	0,025		calculated	calculated from measured outer diameter and thickness
bottom end plug (cm)	16,83				length
top end plug (cm)	25,315				length

Fig. 3.22   Screen shot of parts of the accumulated data to generate the covariance matrices. The data consists of the parameter name and its dimension, its value, uncertainty and (if available) the distribution function. The data bank also states the source of the knowledge w.r.t. [1].

Here the impact of different distribution functions of system parameters is investigated by calculating the k_{eff} values and correlation coefficients for the selected benchmark experiments from series LCT-06, 35 and 62. The most commonly used distributions are the normal N(σ, μ) and uniform U(a, b) distribution functions, where σ and μ are the mean and the standard deviation and a and b the lower and upper bound of the uniform distribution. The modeling approach assumes that within one sample for each experiment and case the fuel composition and fuel rod geometry is identical for all pellet and powder fuel rods respectively. I.e. defining a unit cell for a pellet and a powder fuel rod as shown in 3.23, they are used for all fuel rods in the assembly, in analogy to Scenario A of the
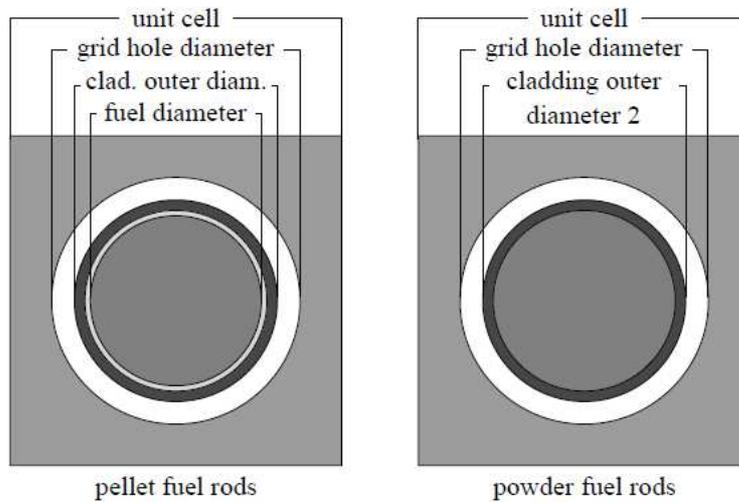
preceding section.

Fig. 3.23 Modeling approach of a unit cell for pellet (left) and powder (right) fuel rods (not to scale).

Applying a Monte-Carlo sampling approach, each uncertain parameter describing the experiment is represented by a distribution function. Using the GRS development SUn-CISTT those parameters are varied according to their distributions and for each benchmark experiment a set of n input samples is generated. For each experiment and case a set of 250 effective neutron multiplication factors k_{eff}^i are calculated. The resulting average values along with the measured experimental values are shown in 3.24. The uncertainties of the experimental values (black in 3.24) are obtained by error propagation and are reported in the experiment documentation. The calculated values and their uncertainties are obtained from the 250 samples for each experiment and case.

The calculated values for experiments from series LCT-06 and 35 are in good agreement with the measured data. Here, the SCALE calculations underestimate the multiplication factor by up to 0.14%, which is a known effect for low enriched uranium setups [26]. For experiments from series LCT-62, however, the multiplication factors are overestimated by up to 0.39% but still agree within a 2σ interval with the measured values. The overestimation might be explained by the presence of powder fuel rods in these experiments. The shape of powder fuel rods might be irregular, making the modeling more difficult and introducing an additional uncertainty, which is difficult to quantify. Furthermore, fig. 3.24 shows that depending on data interpretation the final uncertainty might be underestimated, as
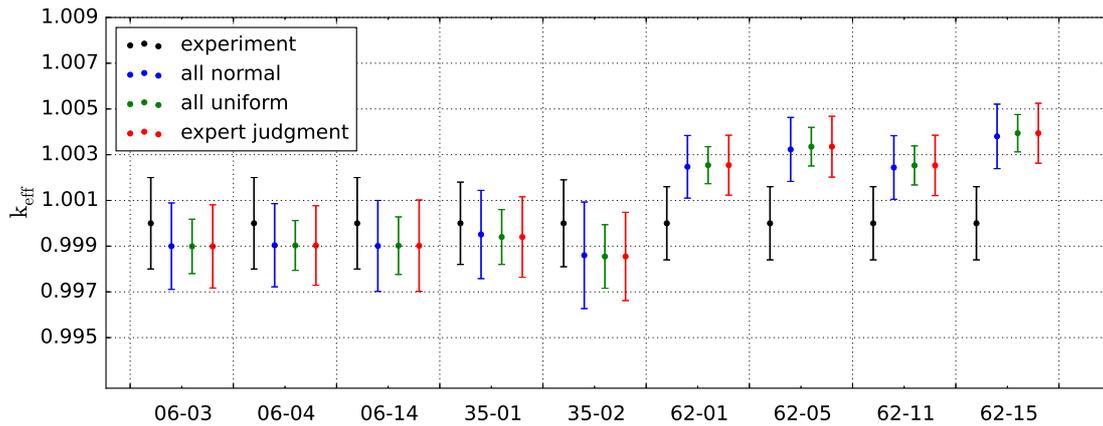
Fig. 3.24 Shown are the k_{eff} values and the 95% CI's for the 9 different experiments assuming all distribution being normal (blue), uniform (green) distributed or a mixture of both (red). Comparing the results to the experimental values (black) shows some disagreement with the experiments from series 62. This might be due to the non-homogeneity of the powder fuel rods due to the manual manufacturing process.

is the case for the data with all uniform distributions (green). Assuming all parameters to be normal distributed leads to a conservative, overestimated uncertainty. Since all three experiment series LCT-06, 35 and 62 share certain components, their calculated $k_{\text{eff}}^{\text{calc}}$ values are correlated. By changing the probability distributions of the parameters, some correlation coefficients change by up to 20%, as e.g. for LCT-06-03 and LCT-35-02 in figure 3.25.

3.4.1 Summary and Discussion

Assuming different distribution functions for system parameters in the Monte-Carlo sampling approach leads to different uncertainties in k_{calc}. While with all parameters uniformly distributed the uncertainty in k_{calc} is smaller compared to the experimental reported one, the case with all parameters normally distributed provides larger uncertainties and thus represents the more conservative approach. In most cases SCALE in combination with ENDF/B-VII nuclear library underestimates the effective neutron multiplication factor. This is also found to be the case for experiments from series LCT-06 and 35. For LCT-62, however the neutron multiplication factors are overestimated possibly due to inhomogeneous distributions of U-235 in the powder fuel rods.

Left matrix (all parameters normal distributed):

	06-03	06-04	06-14	35-01	35-02	62-01	62-05	62-11	62-15
06-03		0.825	0.852	0.788	0.640	0.712	0.684	0.673	0.680
06-04			0.935	0.904	0.762	0.758	0.730	0.743	0.744
06-14				0.916	0.793	0.786	0.765	0.772	0.778
35-01					0.850	0.754	0.723	0.744	0.743
35-02						0.616	0.602	0.611	0.619
62-01							0.961	0.959	0.961
62-05								0.957	0.961
62-11									0.953
62-15									

Middle matrix (all parameters uniform distributed):

	06-03	06-04	06-14	35-01	35-02	62-01	62-05	62-11	62-15
06-03		0.828	0.858	0.818	0.722	0.653	0.672	0.653	0.664
06-04			0.907	0.852	0.747	0.712	0.734	0.702	0.707
06-14				0.909	0.795	0.712	0.740	0.716	0.710
35-01					0.833	0.698	0.734	0.696	0.684
35-02						0.602	0.615	0.631	0.597
62-01							0.902	0.906	0.899
62-05								0.900	0.897
62-11									0.899
62-15									

Right matrix (mixture of normal and uniform distributions – expert judgment):

	06-03	06-04	06-14	35-01	35-02	62-01	62-05	62-11	62-15
06-03		0.817	0.841	0.834	0.801	0.692	0.679	0.664	0.667
06-04			0.936	0.920	0.882	0.756	0.755	0.733	0.752
06-14				0.946	0.910	0.768	0.772	0.744	0.761
35-01					0.932	0.753	0.754	0.726	0.737
35-02						0.715	0.707	0.695	0.705
62-01							0.969	0.953	0.951
62-05								0.953	0.951
62-11									0.949
62-15									

Color scale: 1.00, 0.95, 0.90, 0.85, 0.80, 0.75, 0.70, 0.65, 0.60, 0.55, 0.50

Fig. 3.25 Color coded correlation matrices. Left: all parameters are normal distributed, middle: all parameters are uniform distributed, right: a mixture of normal and uniform distributions referred to as *expert judgment*.

Varying the distribution functions of system parameters also alters the correlation coefficients between the calculated k_{eff}. Some of the correlation coefficients change by up to 20% for the three cases and benchmark experiments considered. However, the most crucial and also time consuming step was the derivation of statistical data from the experimental description. A detailed analysis of the primary experimental documentation was necessary, since the information provided in [1] was not always sufficient.

3.5 Example 3: Shared Components But No Correlation Coefficients

In the following we want to show, that shared experimental setups or components do not necessarily lead to statistical significant correlation coefficients, if these components play only a minor role for the model dependent k_{eff} uncertainty. While the work described above considered water moderated arrays of fuel rods, in the following we examine water reflected spheres of low concentrated plutonium nitrate solution with a thermal neutron spectrum. In the ICSBEP these experiments have the identifier PU-SOL-THERM (PST). We analyze the experiment series PST-03 to -06, and -20 and -21 and determine their correlation matrix. The experiments include different sizes of spheres, different wt.%^{240}Pu and different plutonium nitrate concentrations. One series consists of one size of spheres and several experiments have the same plutonium content of ^{240}Pu. Therefore it is plausible that these experiments are not statistically independent data sets and that correlations arise. A total of 43 experiments from 6 experimental series (PST-03 to -06, and -20 and -21) was analyzed. All experiments describe plutonium nitrate in aqueous solution, contained in metal spheres and are slightly under-moderated. Series PST-03, 04,

Upper (LCT06):

	conU234	conU235	fuelWeight	fuelDiam	fuelHight	cladOutDiam	cladThick	gridHoleDiam	critHight
06-03	-0.076	0.817	0.105	0.045	-0.125	-0.301	0.040	0.059	0.437
06-04	-0.104	0.907	0.130	0.064	-0.191	-0.209	0.061	0.018	0.086
06-14	-0.080	0.964	0.164	0.092	-0.198	-0.088	0.032	0.007	0.009
06-03	-0.211	0.818	0.119	0.064	-0.074	-0.234	-0.006	-0.042	0.476
06-04	-0.166	0.900	0.084	0.093	-0.147	-0.261	0.005	-0.082	0.332
06-14	-0.172	0.952	0.138	0.074	-0.160	-0.086	-0.025	-0.076	-0.044
06-03	-0.170	0.832	0.116	-0.016	-0.113	-0.174	0.129	0.045	0.474
06-04	-0.160	0.898	0.129	0.063	-0.154	-0.115	0.110	-0.033	0.132
06-14	-0.149	0.946	0.180	0.050	-0.181	-0.007	0.115	0.003	-0.078

Middle (-35):

	conU234	conU235	fuelWeight	fuelDiam	fuelHeight	cladOutDiam	cladThick	gridHoleDiam	critHeight	conBTot	conB10
35-01	-0.087	0.927	0.120	0.066	-0.190	-0.220	0.062	0.008	-0.077	-0.084	-0.136
35-02	-0.083	0.879	0.151	0.090	-0.222	-0.203	0.029	0.052	-0.009	-0.239	-0.217
35-01	-0.176	0.891	0.080	0.083	-0.137	-0.180	-0.013	-0.073	-0.014	-0.271	-0.273
35-02	-0.176	0.891	0.080	0.083	-0.137	-0.180	-0.013	-0.073	-0.014	-0.271	-0.273
35-01	-0.177	0.899	0.111	-0.013	-0.178	-0.044	0.146	-0.008	0.101	-0.272	-0.176
35-02	-0.158	0.790	0.110	0.024	-0.074	-0.014	0.134	-0.024	-0.052	-0.415	-0.372

Lower (-62):

	con1U234	con1U235	fuelWeight1	con2U234	con2U235	fuelWeight2	fuelDiam1	fuelHeight1	fuelHeight2	cladOutDiam1	cladThick1	cladOutDiam2	cladThick2	gridHoleDiam	critHeight	conB10	conBTot	GapWidth	PlateThick
62-01	-0.064	0.742	0.163	-0.027	0.488	0.227	0.118	-0.127	-0.117	-0.121	0.014	-0.020	-0.039	-0.046	-0.021	0.000	0.000	0.000	0.000
62-05	-0.063	0.746	0.173	-0.031	0.487	0.220	0.106	-0.139	-0.112	-0.100	0.024	-0.018	-0.045	-0.069	0.106	0.000	0.000	0.063	-0.046
62-11	-0.056	0.714	0.179	-0.033	0.516	0.204	0.086	-0.134	-0.118	-0.112	0.006	-0.043	-0.048	-0.057	0.038	-0.105	-0.004	0.075	-0.007
62-15	-0.076	0.739	0.143	-0.025	0.485	0.210	0.108	-0.158	-0.128	-0.091	0.036	-0.028	-0.069	-0.088	0.058	-0.006	-0.005	0.038	0.051
62-01	-0.122	0.755	0.099	-0.072	0.451	0.224	0.126	-0.170	-0.240	-0.129	0.035	-0.138	0.024	-0.119	0.062	0.000	0.000	0.000	0.000
62-05	-0.122	0.732	0.086	-0.087	0.475	0.236	0.077	-0.182	-0.232	-0.130	0.045	-0.152	0.026	-0.118	0.094	0.000	0.000	0.164	0.020
62-11	-0.097	0.744	0.082	-0.071	0.449	0.238	0.125	-0.183	-0.255	-0.117	0.029	-0.169	0.047	-0.104	0.050	0.036	0.098	0.129	-0.044
62-15	-0.118	0.744	0.073	-0.097	0.438	0.261	0.086	-0.208	-0.237	-0.124	0.017	-0.160	0.063	-0.086	0.006	-0.005	0.038	0.188	-0.029
62-01	-0.072	0.717	0.050	-0.108	0.429	0.265	0.001	-0.167	-0.213	-0.070	0.136	-0.062	0.021	-0.018	0.032	0.000	0.000	0.000	0.000
62-05	-0.047	0.743	0.052	-0.068	0.393	0.232	-0.017	-0.196	-0.229	-0.084	0.132	-0.049	-0.032	-0.014	0.123	0.000	0.000	0.124	-0.029
62-11	-0.086	0.711	0.067	-0.096	0.416	0.237	-0.016	-0.184	-0.250	-0.080	0.143	-0.038	-0.014	-0.047	0.050	0.031	-0.063	0.117	0.012
62-15	-0.036	0.704	0.063	-0.096	0.418	0.232	0.007	-0.198	-0.255	-0.091	0.132	-0.043	0.023	-0.018	0.021	0.041	-0.064	0.087	0.037

Fig. 3.26 Correlation coefficients for the input parameters and the k_{eff} uncertainties for used experiments from series LCT06 (upper), -35 (middle), and -62 (lower). The biggest impact stems for all experiments from the ^{235}U concentration.

05 and 06 have some dissolved iron as impurity in the solution. A list of all experiments can be found in tab. 3.5 showing the 6 calculated series, the considered experiments, the diameter of the spheres in inch[1], the experimental uncertainty, and if the latter was calculated by the evaluator or assumed from similar experiments.

---

[1] The diameters are given in inch, since the original literature is written in United State Customary unit system.

Tab. 3.5   Analyzed experiments and used experimental uncertainty.

Experimental series	Exp.	# Exp.	Color in figures	Diameter of sphere	Experimetal uncertainty
PST-03	01-08	8	black	13"	0.0047 assumed worst case
PST-04	01-13	13	red	14"	0.0047 assumed worst case
PST-05	01-09	9	green	14"	0.0047 assumed worst case
PST-06	01-03	3	blue	15"	0.0035 calculated
PST-20	10-15	6	purple	14"	0.0047 calculated
PST-21	07-10	4	cyan	15.2"	0.0025-0.0044 calculated

Almost all spheres consist of stainless steel. Only experiments PST-03-07 and 08 use an aluminum sphere. For series PST-21 a simplified model is assumed, which does not comprise the metal sphere and is corrected to account for the implications of this modification. Experimentally, almost all spheres are submerged in a rectangular water tank with at least 30 cm of surrounding water. Since from a modeling perspective 30 cm of water reflector is equivalent to an infinite water reflector, a spherical approximation of this water is a valid approximation. Only experiments PST-07, 08 and 09 from series 21 are bare spheres without any reflector. These simplifications and the accompanying compensations allow a very simple, spherical symmetrical computational model of the experiments: the spherical metal tank with the homogeneous plutonium nitrate solution and a surrounding water sphere of 30 cm. Experiments 14 and 15 of series PST-20 have an additional cadmium coating of 0.03 inch (0.762 mm) on the sphere. The detailed descriptions of the experiments in reference [1] include all assumptions and simplifications.

The experiments are modeled with the criticality code sequence CSAS5 of the code packet SCALE 6.1.2 using continuous energy cross-section library ce_v7_endf based on the ENDF/B-VII library. In the CSAS5 calculations 10,000 neutrons are followed, the first 100 generations are skipped and the calculation is stopped, when the Monte-Carlo precision drops below 10^{-5}, a value typically reached after 430 generations.

For the nominal cases, the mass number densities of the solutions, spheres and the surrounding water are taken directly form the experimental description [1]. For plutonium they are given for all isotopes; for nitrogen, hydrogen, oxygen, iron, chrome, nickel, manganese, aluminum and cadmium the natural abundances are used.
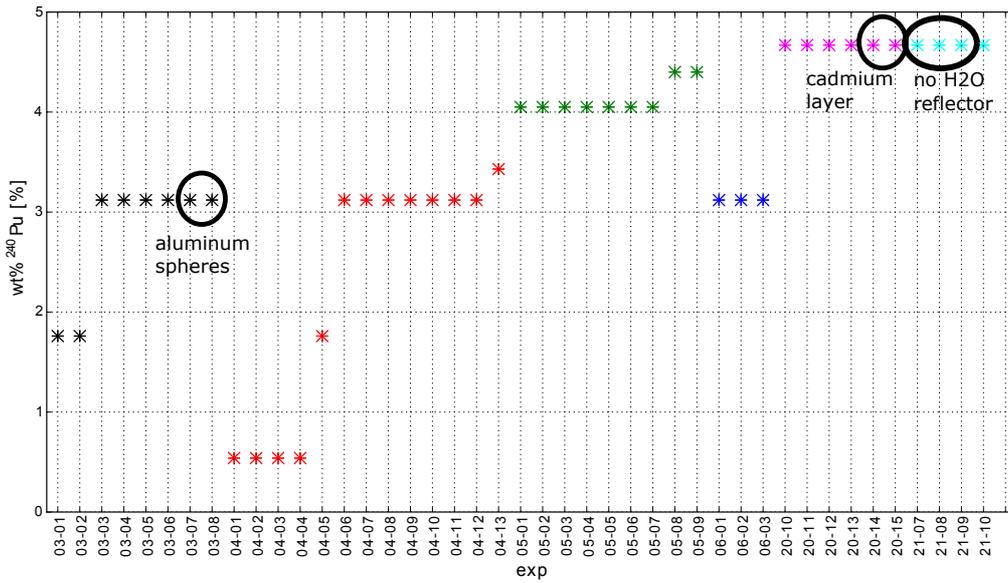
Fig. 3.27 wt.-%^{240}Pu of all considered cases. The black circles indicate the aluminum spheres in series 03, the cadmium layer in series 20 and the lack of H$_2$O reflector in series 21.

Fig. 3.27 shows the wt.-% of ^{240}Pu for all experiments, for different series in different colors. Also the two cases with aluminum sphere, the two cases with additional cadmium layers on the outside of the sphere and the three experiments without water reflector are highlighted. For further details we refer to [31].

series	003								004												
experiment	001	002	003	004	005	006	007	008	001	002	003	004	005	006	007	008	009	010	011	012	013
volume																					
thickWall																					
wt% ^{240}Pu																					

series	005									006			020						021			
experiment	001	002	003	004	005	006	007	008	009	001	002	003	010	011	012	013	014	015	007	008	009	010
volume																						
thickWall																						
wt% ^{240}Pu																						
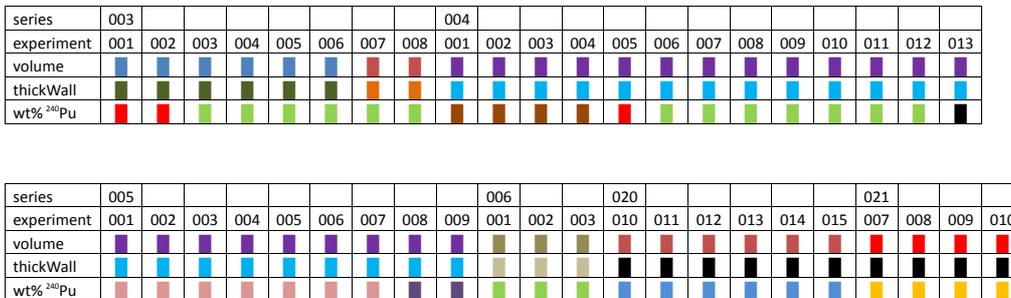
Fig. 3.28 Correlation of varied system parameters. The colors indicate, for which experiments the corresponding parameter is mutually varied. Identical color means identical model parameters.

The experiments are correlated, since they share certain system parameters, which are afflicted by experimental uncertainties. In these geometrically rather simple experiments, the shared system parameters are the volume of the sphere, the thickness of the wall and the wt.-% ^{240}Pu. It is assumed that experiments with the same values are correla-

ted via these parameters. Since the experimental descriptions do not describe how the solutions are mixed from their individual components, all used densities (ρ_{Pu}, ρ_{NO_3}, ρ_{Fe}, ρ_{total}) are assumed to be independent for each experiment. The matrix in fig. 3.28 shows the modeling assumptions for the correlations of varied system parameters. Thereby the parameters are varied mutually for all experiments with the same color box.

We choose 250 samples for each experiment and normal distribution functions for all parameters. The uncertain experimental parameters are listed in tab. 3.6 with the standard deviation of the distribution for each series. The amount of ^{239}Pu (filled up to 100 % Pu), the number densities of the solution and the sphere radii (Pu nitrate, metal, water) for the calculation input are deduced for each sample.

Tab. 3.6   Uncertain experimental parameters.

Uncertain experimental parameters	Variable	Uncertainties	
		series 03, 04, 05, 06	series 20, 21
Total density	ρ_{total}	0.03 %	0.4 %
Pu density	ρ_{Pu}	1.0 %	1.0 %
Fe density in solution	ρ_{Fe}	1.4 %	-
Nitrate density	ρ_{NO_3}	0.6 %	-
Acid molarity	N_a	-	1.0 %
Weight % ^{238}Pu	wt.-% ^{238}Pu	-	16.67 %
Weight % ^{240}Pu	wt.-% ^{240}Pu	7.0 %	0.75 %
Weight % ^{241}Pu	wt.-% ^{241}Pu	-	1.93 %
Weight % ^{242}Pu	wt.-% ^{242}Pu	-	11.11 %
Volume	V	0.3 %	0.25 %
Wall thickness	r_{wall}	10.0 %	-
Temperature H$_2$O	T_{H_2O}	0.09 %	0.09 %

Fig. 3.29 shows the calculated and experimental k_{eff} values for all 43 experiments and the corresponding 1σ deviation. The black crosses show the experimental values $k_{\text{eff}}^{\text{exp}} = 1$, the error bars represent the experimental uncertainties due to the uncertainties of system parameters (see tab. 3.5), which were combined by the evaluators via error propagation. The blue crosses indicate the nominal calculations. Also shown are the mean values and
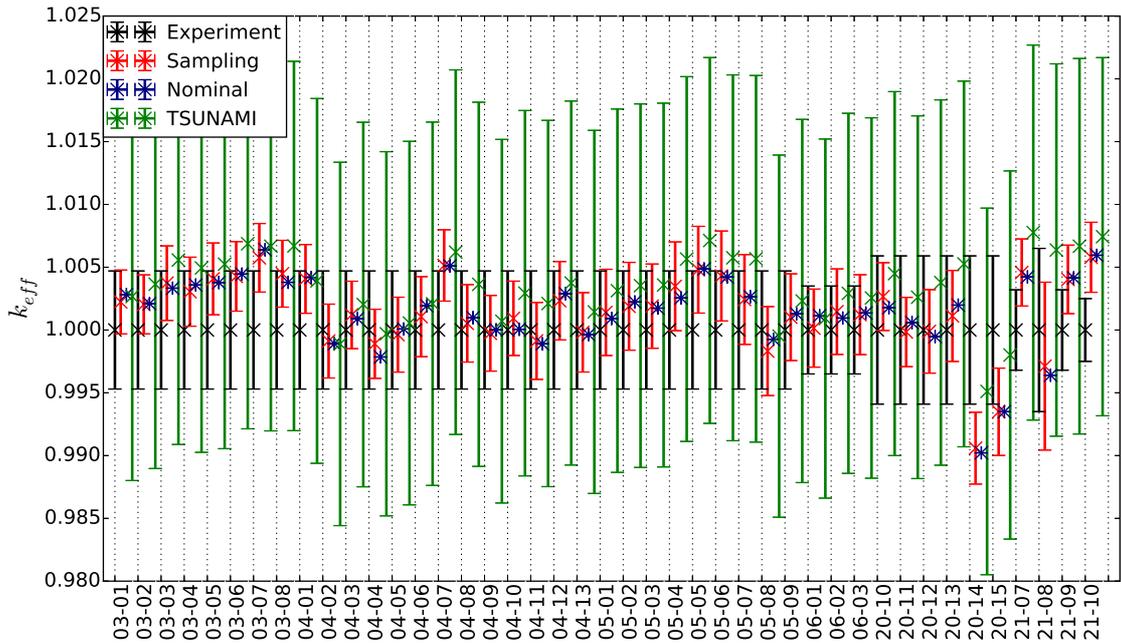
Fig. 3.29 Range of the calculated k_{eff} values and the experimental ones taken from [1] (black). The largest 1-σ uncertainties stem from the uncertainty of the nuclear data, calculated with TSUANMI (green). Note, that most results for sampled (red) and nominal (blue) are in agreement with the experimental values within the 1-σ error bars. The remaining agree within th 2-σ interval.

standard deviations of the sampling calculations due to the variation of system parameters (red) and of nuclear data (green, TSUNAMI).

Compared to the other experiments the two experiments with cadmium coatings of the spheres (PST-20-14 and 15) deviate significantly towards lower values ($\Delta k_{\text{eff}} \approx -0.0112$, and -0.0080 respectively). The hypothesis for this effect is, that KENO-Va overestimates the influence of cadmium, possibly due to nuclear cross sections. For a deeper discussion of the Cadmium in these experiments we refer again to [31]. However, the two experiments are used in the following with the original assumed 0.03 inch of cadmium coating.

The calculated k_{eff} values and their standard deviations are included in fig. 3.29 in red. The mean values agree very well with the nominal values. The standard deviations are in a comparable range to the experimental uncertainties, but for series PST-03, 04, 05 and 20 they are 30-50 % smaller. This can be attributed to the fact, that for these series, the given experimental uncertainty is not calculated directly, but assumed from calculations of

other very similar experiments. For series PST-06 and 21 we find an excellent agreement of the standard deviations and experimental uncertainties.
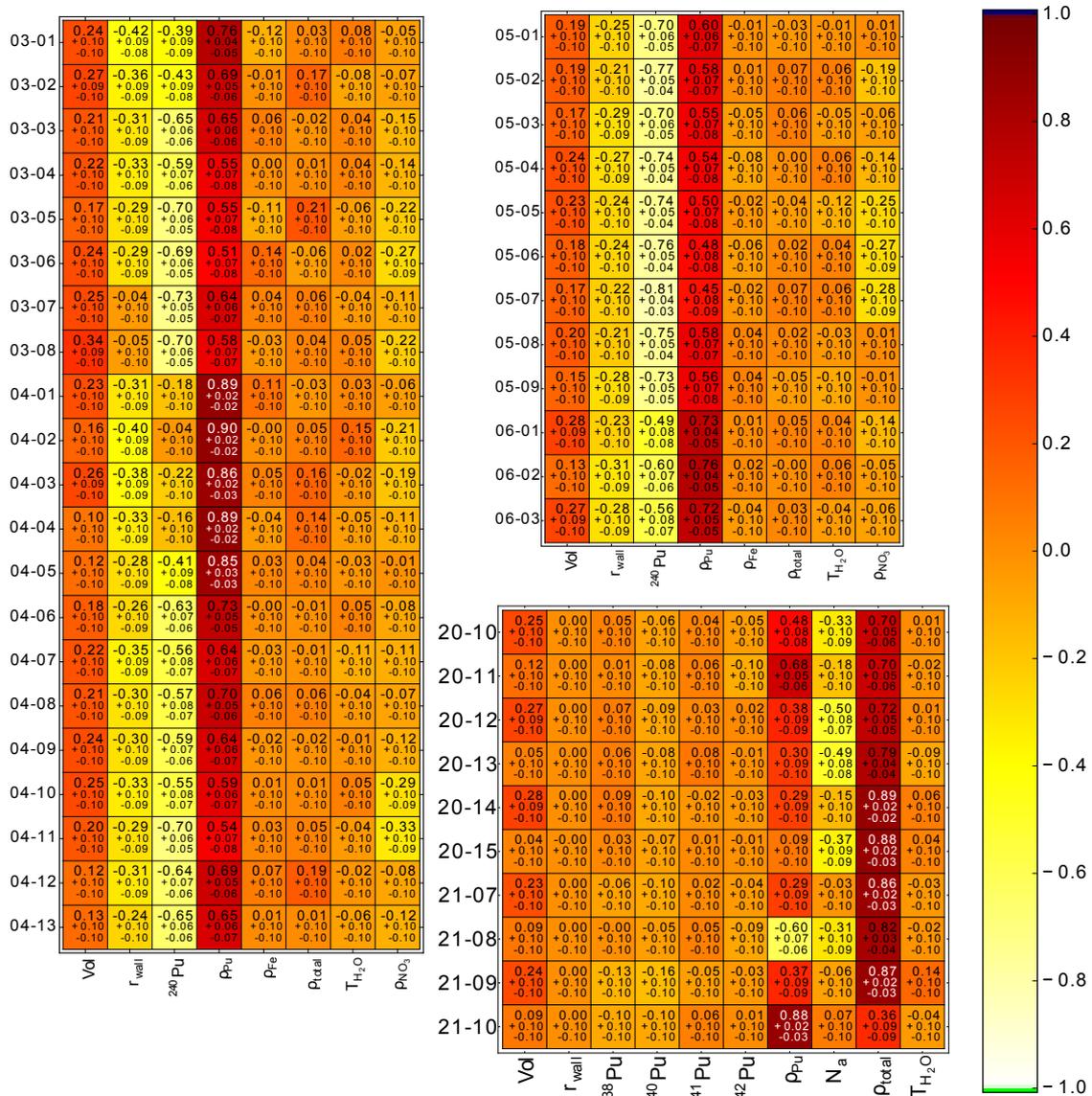
Fig. 3.30 Sensitivities of k_{eff} on uncertain parameters for all analyzed experiments. Numerical values shown for mean values and the 95% confidence levels.

For all experiments the correlation coefficients were calculated between the varied parameters and k_{eff}, shown in fig. 3.30. This gives a measure for the influence of the variation of each uncertain parameter on the uncertainty of k_{eff} and can demonstrate the leading effects. Note that this is not the sensitivity of k_{eff} on the uncertain parameters, but the sensitivity of k_{eff} on the actual variation of the uncertain parameters.

For the first four series (PST-03, to -06), the two leading effects are a negative correlation between k_{eff} and wt.-%^{240}Pu, and a positive correlation with the Pu density ρ_{Pu}. The first relation can be explained by the reduction of the most reactive plutonium isotope ^{239}Pu. The second by an increase of plutonium atoms available for fission. The k_{eff}-decreasing effect of the small decrease of H/Pu by the increase of ρ_{Pu} in these under-moderated systems can be neglected. A small positive correlation exists to the sphere volume V.

All these correlations are obvious: The more material is present, the higher is k_{eff}. Also notably is that the sensitivity to wt.-%^{240}Pu increases with wt.-%^{240}Pu itself since the absolute variation increases. The small negative correlation with the wall thickness r_{wall} can be understood by the neutron absorption of stainless steel. Therefore its value is not significantly different from zero for the two experiments with aluminum sphere PST-03-07 and 08. The uncertainty of the density of the impurity iron ρ_{Fe}, of the total density ρ_{total} and of the temperature T_{H_2O} have almost no significant effect on k_{eff}. For the total solution density ρ_{total} this is certainly attributed to the very small given uncertainty of only 0.03 % for these experiments.

For the second set of experiments (PST-20 and -21) the situation is slightly different. Here the leading effect is the total solution density ρ_{total}, which has a 13 times higher uncertainty of 0.4 % leading to a strong positive correlation. Additionally the given uncertainty of wt.-%^{240}Pu in this second set is a factor of 100 smaller than in the first set, so that its influence on k_{eff} disappears almost completely. The next effect is a negative correlation with the acid molarity N_a. An increase of N_a leads to an increase of ρ_{NO_3} and a decrease of ρ_{H_2O}. ρ_{NO_3} increases the number density of ^{14}N having a considerable neutron absorbing effect, ρ_{H_2O} drives the moderation ratio away from its optimum value, both explaining the negative impact on k_{eff}. The mostly positive correlation with ρ_{Pu} is evident due to the same effect as for the first set of experiments.

The fig. 3.31 shows the correalation coefficients for all 43 experiments due to the variation of system parameters In general, the correlation values range from slightly negative values to 0.7, but most correlation coefficients are in the range of [-0.1,0.3]. Since the error of cor is in the range of 0.1 for values around $cor =0$ for the used 250 samples [32], most correlations between experiments can be considered statistically not or only slightly significant. Higher correlation coefficients can be found within the experimental series PST-03 (experiments 03 to 08), PST-04 (experiments 06 to 12), PST-05 (experiments 01
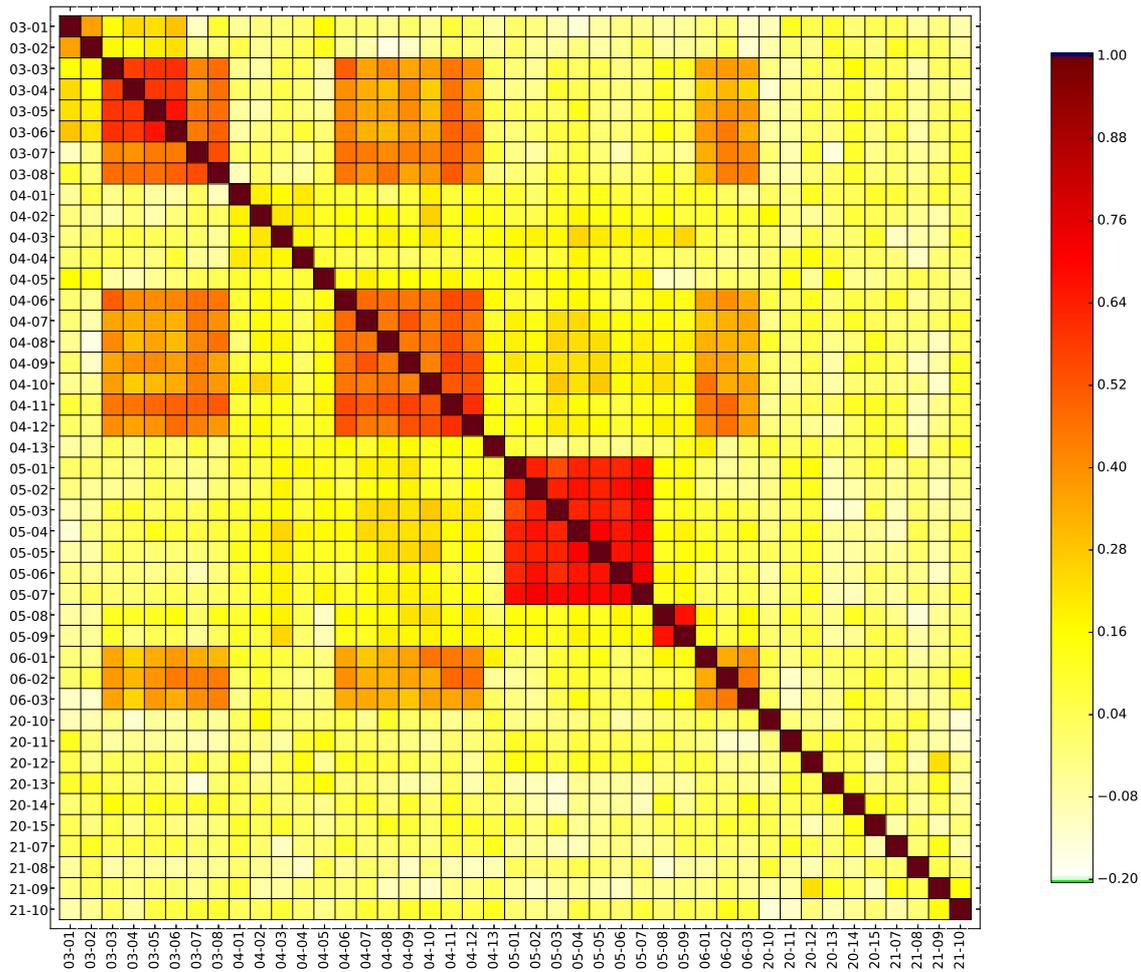
Fig. 3.31 Correlation matrix of k_{eff} between all experiments due to the partly mutual variation of uncertain system parameters. For most of the experiments a statistical significant correlation coefficient could not be found. Only some experiments of series 03,04,05, and 06 show some correlation coefficients.

to 07 and 08, 09) and PST-06 (experiments 01 to 03). Further, three blocks of higher correlation coefficients between experimental series can be identified: Experiments PST-03-03 to -08 with PST-04-06 to -12 and experiments PST-06-01 to -03, and PST-04-06 to -12 with PST-06-01 to -03.

All cases of series PST-20 and 21, PST-04-01 to -05 and -13 are uncorrelated to the others. This can be explained, since the sensitivies of k_{eff} on the mutual varied parameters are much smaller than the ones on the individually varied parameter.

3.5.1   Summary and Discussion

Comparing the results shown in fig. 3.31 with fig. 3.28 one can see, that almost all blocks of higher correlations are due to the same wt.-%^{240}Pu. This is the only one of the two leading effects of the sensitivity of k_{eff}, which is varied mutually: PST-03-01 and -02, PST-03-03 to -08, PST-04-06 to -12, PST-05-01 to -07, PST-05-08 and -09, PST-06-01 to -03, and PST-03-03 to PST-03-08 with PST-04-06 to -12 and with PST-06-01 to -03.

The results can be compared to available data of the *Database for the International Criticality Safety Benchmark Evaluation Project* (DICE [33]). DICE is still in the development phase and also subject to possible data entry errors and omissions. However, it provides some information on correlations on a vast number of experimental series described in the ICSBEP Handbook [1]. The relevant data for our analysis is shown in tab. 3.7. There is no data available neither for the correlation of PST-21 nor for results on the case level details. A '+'-sign indicates strong correlations between experiments when one or several uncertain benchmark parameters are correlated, which are major contributors to the overall benchmark k_{eff} uncertainty. A '(+)'-symbol indicates a 100% correlation.

The currently available data in DICE does only partly agree with the findings of our more detailed analysis (fig. 3.31). On one side, we also found statistically significant correlation coefficients between series for experiments PST-03-03 to -08 with PST-04-06 to -12 and PST-06. However all these cases have correlation coefficients below 0.6.

On the other side, all experiments of PST-20, PST-21 PST-04-01 to -05 and -13 show no statistically significant correlation coefficients with any other investigated experiment, even within the same series.

The only variation introducing a correlation effect between different series is the identical wt.-% ^{240}Pu of the experiments PST-03-03 to -08, PST-04-06 to -12, and PST-06. These results support the data provided by DICE. We found, that for these experiments, the correlation due to shared experimental components are small. We find further, that certain experiments (PST-03-01 and -02, PST-04-01 to -05 and -13, series PST-20 and -21) have no statistical significant correlation coefficient with experiments from any other series. The shared components for these experiments do not introduce correlations due to their small given uncertainty in the experimental description.

Tab. 3.7   Information on possible correlations due to shared experimental components taken from [33]. A '+'-sign indicates strong correlations, a '(+)'-symbol 100% correlation.

	PST-03	PST-04	PST-05	PST-06	PST-20
PST-03	(+)	+	+	+	+
PST-04	+	(+)	+	+	+
PST-05	+	+	(+)	+	+
PST-06	+	+	+	(+)	+
PST-20	+	+	+	+	(+)

This is in contrast to e.g. our findings for lattices of fuel rods described in the subsections before where shared experimental components can introduce large correlation coefficients in the data due to higher sensitivities of k_{eff} to these shared components and their corresponding uncertainties.

The presence of shared components within an experimental series does not necessarily lead to statistically significant values of correlation coefficients. If the shared experimental components are very well known in the sense of comparable small uncertainties or if these components play only a minor role for Δk_{eff} (determined e.g. by means of a sensitivity analysis), their contribution to the correlation coefficient is negligible.

3.6   Summary

The generation of integral experiment covariance matrices of experimental series as published for example in [1] needs two steps: Analyzing the experimental data and calculate the covariance matrices based on it.

Technically calculating covariance or correlation matrices is straight forward. We have shown in section 3.1 a Monte-Carlo sampling based method, orchestrated by SUnCISTT [22]. However, applying Monte-Carlo sampling methods, one needs always to ensure that the results converge. In section 3.3.2 we have shown the effect of an uncertainty (here the $\sigma_{k_{\mathrm{eff}}}^{MC}$ of the Monte-Carlo based neutron transport code KENO.Va) overlapping the experimental uncertainty and thus in a sense hiding correlation effects. We found for our example, that a $\sigma_{k_{\mathrm{eff}}}^{MC} \approx 10^{-4}$ leads to converging correlation coefficients. Another crucial

step for Monte-Carlo based analysis in general is the sample size. We found that our results are always converged for sample sizes larger than 250. If the data is close to being full (anti-)correlated, meaning a correlation coefficient of almost (-)1, the results converge faster, sometimes even after 50 samples.

The more difficult task is to define the modeling based on the available experimental data. We have shown in section 3.3.1 that based on the data provided in the UACSA Phase IV benchmark description of the WPNCS at the OECD-NEA, several modeling approaches can be chosen, leading to different sensitivity profiles of the resulting k_{eff} values and correlation coefficients varying from 0 to 1. Projecting these finding to the publicly available experimental data evaluated in [1], the freedom of modeling choice seems to be difficult to circumvent. To create reliable covariance data based on the data in [1] further work has to be done. This would include a detailed analysis of the original data and extract further information. For our example of lattices of fuel rods one could narrow down the 9 scenarios if one would know e.g. if a fuel rod remained in its grid position for the whole series or if it was randomly placed for each individual experiment.

Another obstacle is the extraction of statistical data like uncertainties and distribution functions of parameters, discussed in section 3.4. A lot of time was spent in deriving parameters from the original literature and document it in a data base. The impact of varying distribution functions (all normal or uniform distributed and a mixture of both) has been discussed.

In conclusion, further work has to be conducted to clarify, how to treat gaps in the experimental data on the way of generating reliable covariance matrices. Another huge obstacle in generating reliable covariance matrices based on the full Monte-Carlo sampling is the calculation time. The generation of covariance data for three experiments, LCT39 6, 7, and 8 needed up to 60k CPU-h. Future work needs to lower this time significantly.

4   Use of Covariance Matrices for Validation purposes

In the following we use and explain the Bayesian updating of an application case k_{eff} using benchmark experiments. We demonstrate the use of formulas 2.21 and 2.22 by discussing a Toy Model, based on the one suggested in [11]. We than use the integral covariance matrices generated in the preceding chapter with the corresponding covariance matrices due to nuclear data to predict some example application cases k_{eff}'s.

The covariances or correlation coefficients between benchmark experiments and application cases due to manufacturing uncertainties are strongly connected with the definition and interpretation of the uncertainty afflicted parameters. On the other side the experiments are correlated due to uncertainties of the nuclear data of the involved atomic processes. These are mainly dependent on the material composition and can be assumed independent from system parameter uncertainties. Therefore the two uncertainties will be calculated separately. For this work the SCALE 6.1.2 Sequence TSUNAMI-3D-K5 was used to calculate the sensitivities of k_{eff} on the participating nuclear processes and TSUNAMI-IP for the determination of the Pearson correlation coefficients, the so called c_k-values, from which the covariances can be derived.

4.1   Example 1: A Toy Model Analysis

4.1.1   Reproducing "True" Values

In the following the described method will be tested against the question, how good the model can predict a bias, or in other words: How good is the posterior k_{eff}-distribution compared to the real (but unknown) value $k_{\text{eff}}^{\text{real}}$. We assume that we know all data (nuclear interactions and geometrical parameters) of a Toy Model exactly and compare the calculated result for k_{eff} with results from distorted data updated with several benchmark experiments.

Assuming a simple Toy Model based on the one proposed in [11], the k_{eff} of a nuclear system is described by

$$k_{\text{TM}}(\vec{\alpha}, \vec{x}) = \frac{(\alpha_1 \alpha_4 x_1)}{(\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3)} \qquad (4.26)$$

with $\vec{\alpha}$ being the vector of relevant nuclear interactions and $\vec{x}$ the vector of the system describing parameters. Assuming the real vectors to be $\vec{\alpha}_r = $(1,2,3,4) and $\vec{x}_r = $(3 ,2 ,1.9)

leads to k_{TM}^r =0.94488. In the following, it will be investigated, to which extend the resulting k_{TM} from the Bayesian updating process using distorted data matches k_{TM}^r. The distorted data might for example represent a cross section library in a given transport code or measured system parameters. These values are always accompanied by systematic and non-systematic uncertainties (or errors), hiding the real values.

For the distorted vectors we assume normal distributed 1-σ errors of 5 % of the real value, leading to the two (randomly generated) vectors $\vec{\alpha}_{\mathrm{dis}}$ =(1.0811, 1.8635, 2.8105, 3.9698) and $\vec{x}_{\mathrm{dis}}$ =(1.0811, 1.8635, 2.8105, 3.9698). Possible combination of theses vectors and the resulting k_{TM}'s are shown in table 4.8.

Tab. 4.8   Resulting k_{TM} and deviations from the real value for different sets of $\vec{\alpha}$ and $\vec{x}$

	$(\vec{\alpha}_r,\vec{x}_r)$	$(\vec{\alpha}_r,\vec{x}_{\mathrm{dis}})$	$(\vec{\alpha}_{\mathrm{dis}},\vec{x}_r)$	$(\vec{\alpha}_{\mathrm{dis}},\vec{x}_{\mathrm{dis}})$
k_{TM}	0.9448	0.94985	1.04590	1.051051
$k_{\mathrm{TM}}^r - k_{\mathrm{TM}}$	0	-0.00497	-0.10102	-0.10617

The set $(\vec{\alpha}_{\mathrm{dis}},\vec{x}_{\mathrm{dis}})$ leads to the largest deviation of more than 10 % in k_{TM} compared to the real case. The deviation caused by $\vec{\alpha}_{\mathrm{dis}}$ is larger compared to the one introduced by $\vec{x}_{\mathrm{dis}}$ which leads only to a very small deviation in the resulting k_{TM}.

Now we have calculated three biases for the different cases. The realistic combination is the one with both vectors distorted, $\vec{\alpha}_{\mathrm{dis}},\vec{x}_{\mathrm{dis}}$. We continue our analysis with this case. Let's assume we have now 9 benchmark experiments where we can calculate the (biased) k_{eff} values and compare them to the experimental values. Due to lack of knowledge of the true nuclear data, we always take the distorted one. This can be seen as a library used in a neutron transport code, e.g. KENO. The bias b introduced in our model by this is thus b =-0.10102.

In the next step, $k_{\mathrm{TM}}(\vec{\alpha}_{\mathrm{dis}},\vec{x}_{\mathrm{dis}})$ will be updated with additional data to see how good the true value k_{TM}^r can be matched. The data shown in table 4.9 describes 9 additional data sets called BM1 to BM9. We assume, that the data is pairwise uncorrelated. Correcting the k_{TM} values for the bias introduced by the nuclear data of the sampled results for $k_{\mathrm{TM}}^{\mathrm{mean}}(\vec{\alpha}_{\mathrm{dis}}, \vec{x}_{\mathrm{BM}})$ leads to values around 0.94. The covariance matrix due to the nuclear data of the application case and the BM's (Σ_{0B} in equations 2.21 and 2.22) for 1000

Tab. 4.9   Additional data used for the updating process.

	x_1	x_2	x_3	$k_{\exp}$	$k_{\mathrm{TM}}^{\mathrm{mean}}(\vec{\alpha}_{\mathrm{dis}}, \vec{x}_{\mathrm{BM}}) \pm 1\sigma$
BM 1	3.0072	1.82	2.025	0.9396	1.0468 ± 0.01883
BM 2	3.0072	1.9292	1.952	0.9402	1.0460 ± 0.01898
BM 3	3.0072	3.9477	0.6151	0.9409	1.0460 ± 0.01986
BM 4	3.0072	6.065	-0.7895	0.9406	1.0471 ± 0.01919
BM 5	3.0072	8.037	-2.097	0.9403	1.0464 ± 0.01970
BM 6	3.0072	9.8448	-3.297	0.9406	1.0478 ± 0.01992
BM 7	3.0072	15.9819	-7.365	0.9401	1.0460 ± 0.01882
BM 8	3.0072	19.9995	-10.029	0.9403	1.0469 ± 0.01950
BM 9	3.0072	23.9692	-12.662	0.9412	1.0462 ± 0.01959

samples can be translated in a matrix of correlation coefficients shown in figure 4.32. The assumption, that the data describing the BM's is uncorrelated is shown in figure 4.33. The values correspond to the one in the covariance matrix named Σ_B in equations 2.21 and 2.22. Using the eq. 2.21 and 2.22 we can now update the prior knowledge of the application case $k_{\mathrm{TM}}^{\mathrm{Apl\,prior}}=1.0527 \pm 0.00529$. The result for the posterior distribution is $k_{\mathrm{TM}}^{\mathrm{Apl\,post}} = 0.94797 \pm 0.010437$. This result matches the real value of $k_{\mathrm{TM}}^{r}=0.94488$. The prior and posterior distribution functions are plotted in figure 4.34. Fig. 4.35 shows the posterior k_{TM} value on the y-axis plotted versus the number of BM's considered in the updating process. We found for the example discussed here, that already the first experiment pushes the posterior value to the k_{TM}^{r} value. After adding BM3, no significant change for the posterior value can be observed.

4.1.1.1   Summary and Discussion

We have shown that the Bayesian updating method described in section 2.1 as introduced in [7] is capable of predicting the bias of an application case if the biases of the benchmarks are known. If the data used for the updating process consists of experimental data with an additional and unknown bias, e.g. due to human errors or a defect instrument, this bias would naturally affect the posterior distribution. Assuming the BM k_{TM} would include an additional bias of 1%, the example described above would lead to a posterior value of $k_{\mathrm{TM}}^{\mathrm{Apl\,post,bias}}=0.95793 \pm 0.010759$, see figure 4.36. The Bayesian updating process presented can only adjust to an overall bias. If the bias stems not only from geometri-
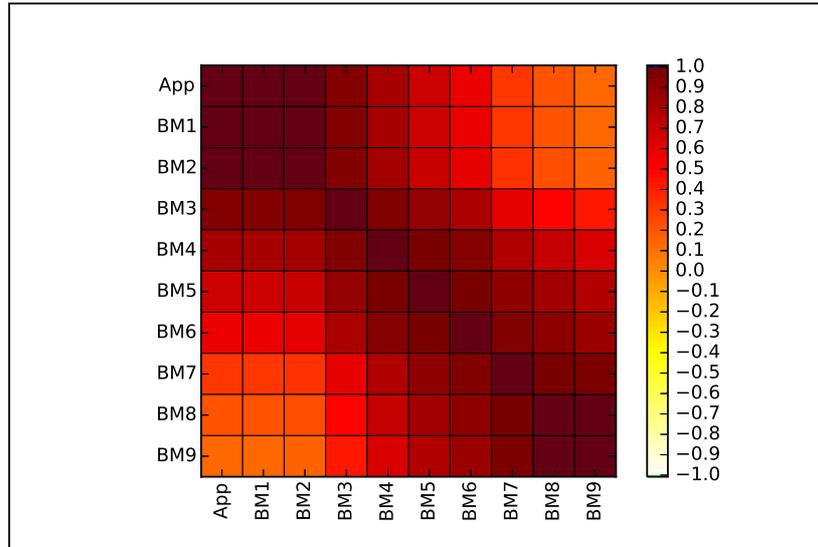
Fig. 4.32   Correlation coefficients due to nuclear data for the application case and BM's. The values indicate a high similarity w.r.t. to the neutron spectrum of the application case and the benchmark experiments with a slow decrease from BM3 to BM9.

cal and nuclear variations but further, e.g. experimental ones, the predictions for the true k_{TM} is affected by this. For the optimal use of this method, the data used for the updating process and especially the biases need to be known in detail.

4.1.2   Influence of Correlated Data on Predicted Values

Similar to the preceding section, we investigate the influence of correlated data on the posterior distribution by discussing a Toy Model. The latter is taken from the UACSA Benchmark Proposal Phase IV [11]. The k-value is defined via equation 4.26 with the values shown in table 4.10. The calculated k_{TM} do always exceed the experimental values. Assuming normal distributed k_{TM} the deviation is in average $\delta k = 0.0121 \pm 0.0059$. The nuclear data vector α_i for $i = 1, ..., 4$ and the corresponding covariance matrix $\Sigma_\alpha = \mathrm{diag}[\sigma^2_{\alpha_i}] = \mathrm{diag}[10^{-4}]$ is also given. From the given data, the necessary matrices and vectors for the Bayesian analysis following equations 2.21 and 2.22 were created. Fig. 4.37 shows the color coded correlation coefficients for the matrices cov_0 (left picture) and $\mathrm{cov}^{\mathrm{sys,B}}$ assuming independent variation of x_1 (right picture). Fig. 4.38 shows color coded matrices for $\mathrm{cov}^{\mathrm{sys,B}}$ with higher correlation coefficients representing more dependencies in the data. The resulting prior and posterior distribution characteristics for the different $\mathrm{cov}^{\mathrm{sys,B}}$ are shown in table 4.11. For comparison, an additional, simple
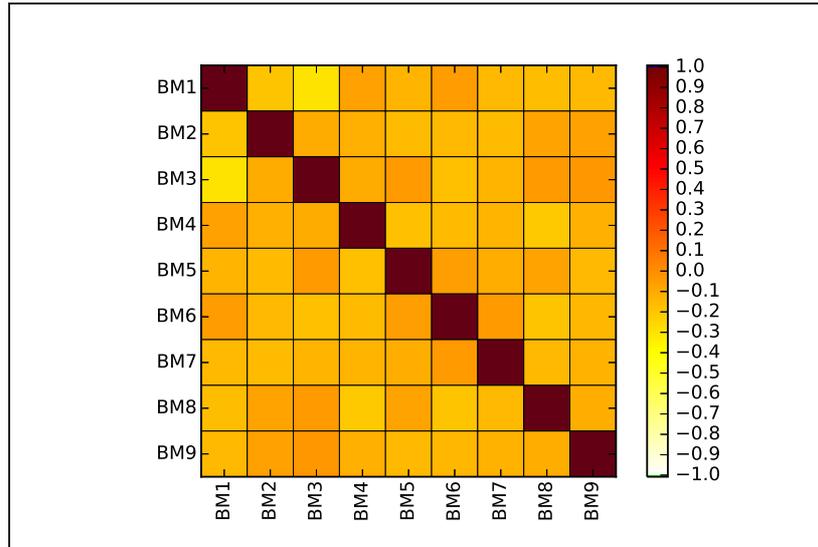
Fig. 4.33   Correlation coefficients due to varied $\vec{x}$ values of the BM's. The values show no higher correlations.

linear prior correction is given, based on the average deviation of the calculated k_{TM}'s. The posterior distribution characteristics show a lower k_{TM} and significantly smaller 1σ deviations than the prior. It follows the tendency that the higher the correlation is, the closer is the k_{TM} estimation of the prior and the larger is the 1σ deviation.

This effect can be seen in fig. 4.39. On the left, the prior and the three different posterior distributions are shown. The solid, dashed, and dotted lines show the projections of the corresponding mean values. The right picture shows the posterior mean and the 1σ uncertainty of the posterior for a stepwise updating of the prior. It shows the direct influence of the different $\mathrm{cov}^{\mathrm{sys,B}}$ on the posterior. Note that the shape of the dashed curves is relatively equal after updating with the third Benchmark. This effect arises from the term $k_{\mathrm{eff}}^{\mathrm{B,exp}} - k_{\mathrm{eff}}^{\mathrm{B,th}}$ in equation 2.21, equal to all posteriors.

It is worth noting, that the Monte-Carlo sampling for 250 samples leads always to the same k_{TM} values (within the standard deviation), only the covariance/correlation matrices vary. Fig. 4.40 shows the matching of the k_{TM} for sampled α and correlated and uncorrelated x with the values given in [11]. The values show a perfect matching for each BM, as one would expect.
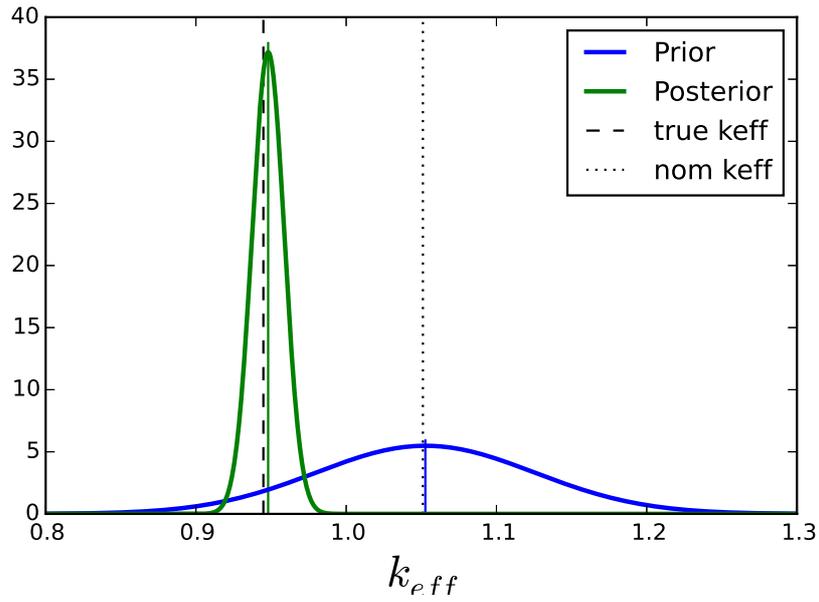
Fig. 4.34 Comparison of prior (blue) and posterior (green) distribution functions and the *true* (dashed line) and nominal k_{eff} value (dotted line). The solid green and blue lines indicate the respective mean values. The posterior distribution function reproduces the *true* value.

Tab. 4.10 Values for the Toy Model analysis, taken from [11].

ID	x_1	σ_1	x_2	σ_2	x_3	σ_3	k_{TM}	k_{exp}
BM1	2.0072	0.05	4.0424	0.05	-0.0746	0.05	1.0174	1.0
BM2	2.0072	0.05	1.9601	0.05	1.9292	0.05	1.0194	1.0
BM3	2.0072	0.05	-0.0506	0.05	3.9477	0.05	1.0177	1.0
BM4	2.0072	0.05	-2.0458	0.05	6.0650	0.05	1.0111	1.0
BM1	2.0072	0.05	4.0424	0.05	-0.0746	0.05	1.0174	1.0
BM5	2.0072	0.05	-3.9905	0.05	8.0370	0.05	1.0086	1.0
BM6	2.0072	0.05	-6.0613	0.05	9.8448	0.05	1.0185	1.0
BM7	2.0072	0.05	-12.0059	0.05	15.9819	0.05	1.0063	1.0
BM8	2.0072	0.05	-16.0923	0.05	19.9995	0.05	1.0066	1.0
BM9	2.0072	0.05	-20.0440	0.05	23.9692	0.05	1.0032	1.0
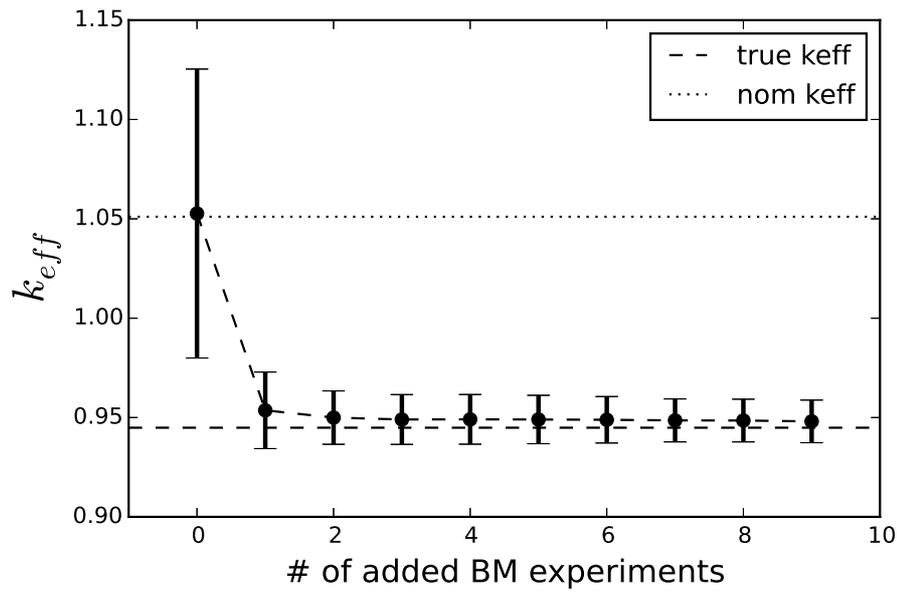
Fig. 4.35   Impact of experiments added to the updating process on the resulting posterior k_{eff} distribution. Shown are the mean values and the 1-σ deviation and the *true* (dashed line) and nominal (dotted line) k_{eff} values. After adding the third experiment, the changes of the mean value and deviations are comparably small. An effect also due to the decreasing correlation coefficients due to α for the application case and BM's 3 to 9.

Tab. 4.11   Table of prior and posterior k_{TM} distribution characteristics for the different covariance matrices

	Prior	Prior - δk	Post lower cor	Post medium cor	Post higher cor
k_{TM}	0.9533	0.9412	0.9424	0.9447	0.9479
1σ	0.0128	0.0766	0.0084	0.0090	0.0107

59
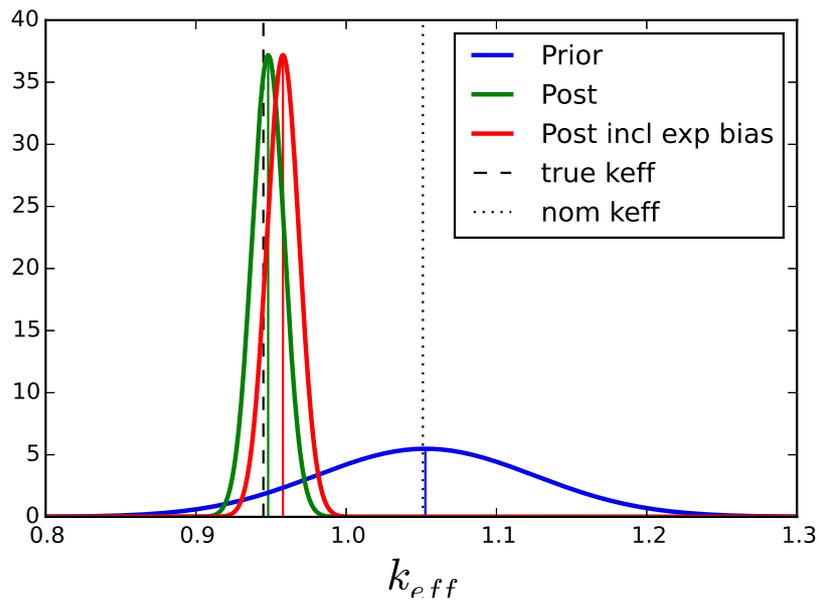
Fig. 4.36 Characteristics of prior (blue) and posterior (green) distributions including an unknown experimental bias (red). The updating process relies strongly on the experimental data and thus can not correct any unknown bias in this data, as shown by the green and red curve.
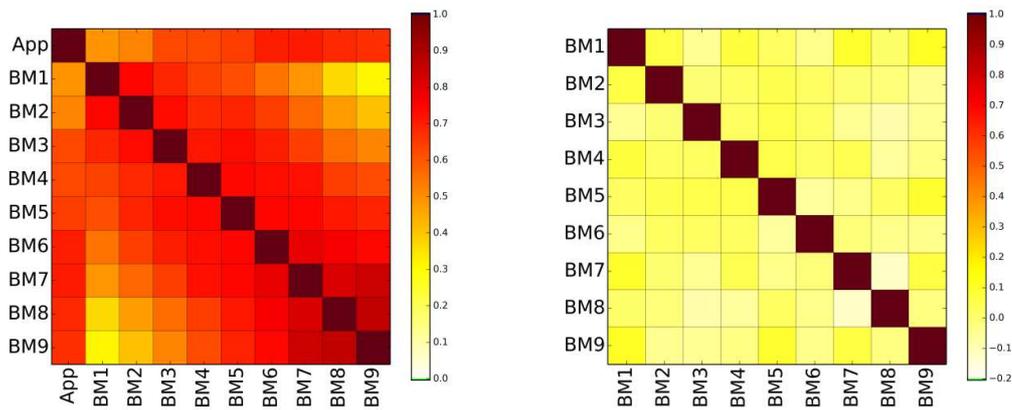
Fig. 4.37 Color coded correlation coefficients for the application case and the 9 Benchmark experiments due to nuclear data variations (left) and correlation coefficients between the benchmark experiments due to variation of system parameters for independent variations of x_1.
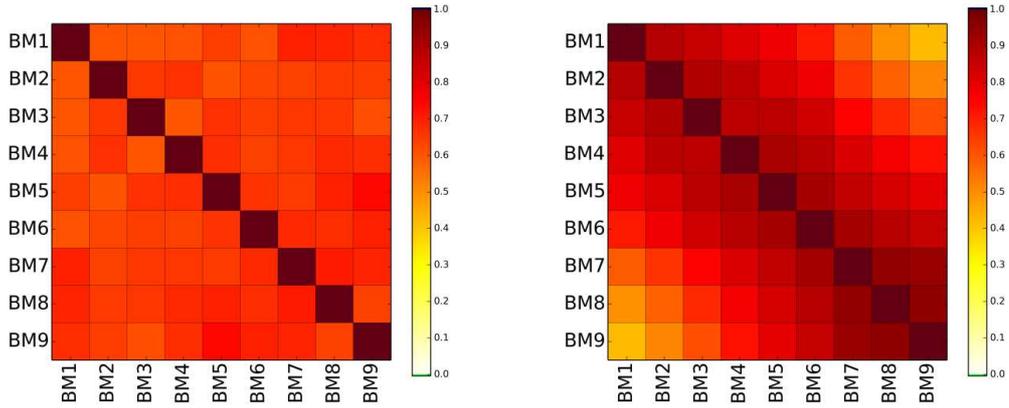
Fig. 4.38 Color coded correlation coefficients for the benchmark experiments due to system parameters for different dependencies of the system parameters. The left part shows a hypothetical constructed matrix, the right one assumes all x_1 to be identical.
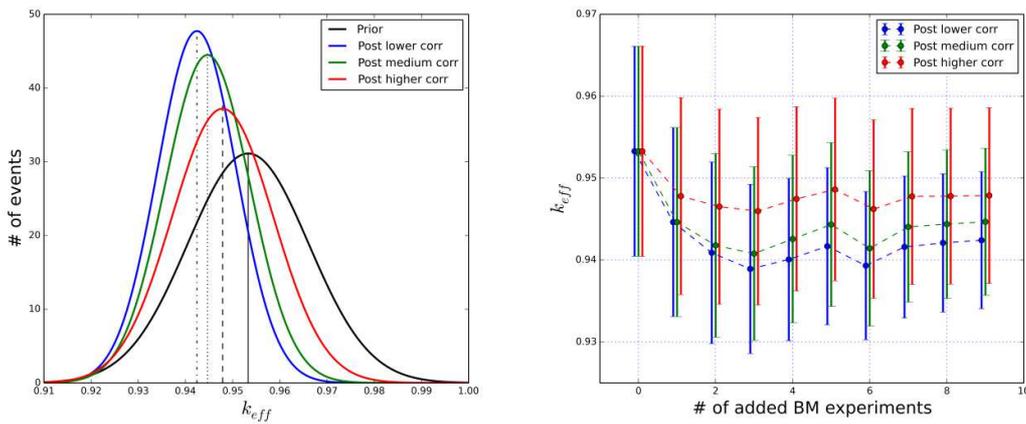
Fig. 4.39 Prior and posterior distribution functions for different covariance matrices (left) and the posterior mean and 1σ uncertainties for a stepwise updating (right).
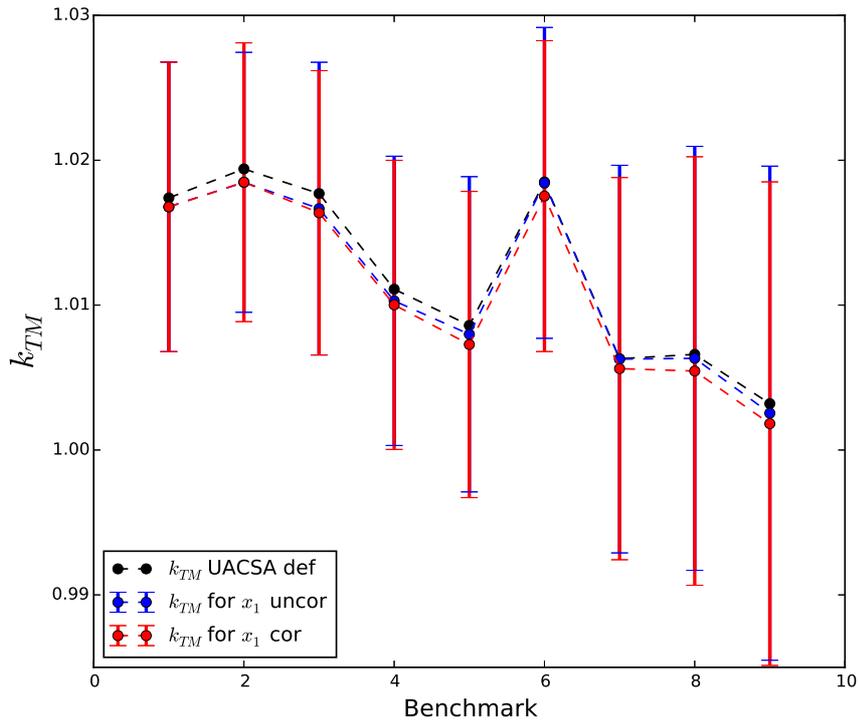
Fig. 4.40 Comparison of k_{TM} from the UACSA benchmark proposal [11] and for sampled α and correlated and uncorrelated x_1 values

4.2 Application using Benchmark Experiments

4.2.1 Example 1: Different Modeling Scenarios

The Bayesian updating procedure will now be applied using some real experimental data evaluated and documented in the ICSBEP handbook [1] and two application cases, as proposed in [11]. The experiments are all water moderated, low enriched uranium fuel rods from the series LEU-COMP-THERM-07 and -39, described in section 3.3. Since all experiments were conducted at the same laboratory, partly using the same equipment, the systematic uncertainties lead to correlated k_{eff} results. The correlation coefficients due to system parameters depend strongly on the model assumptions, as shown in section 3.3.1. The matrix of color coded correlation coefficients due to nuclear data variations is shown in fig. 4.41 including both application cases, App.1 and LCT-79-01. The application case App.1 is a simplified 16×16 UO_2 PWR fuel assembly, fully reflected by water. The data was taken from [11]. The second application case is the experiment LEU-
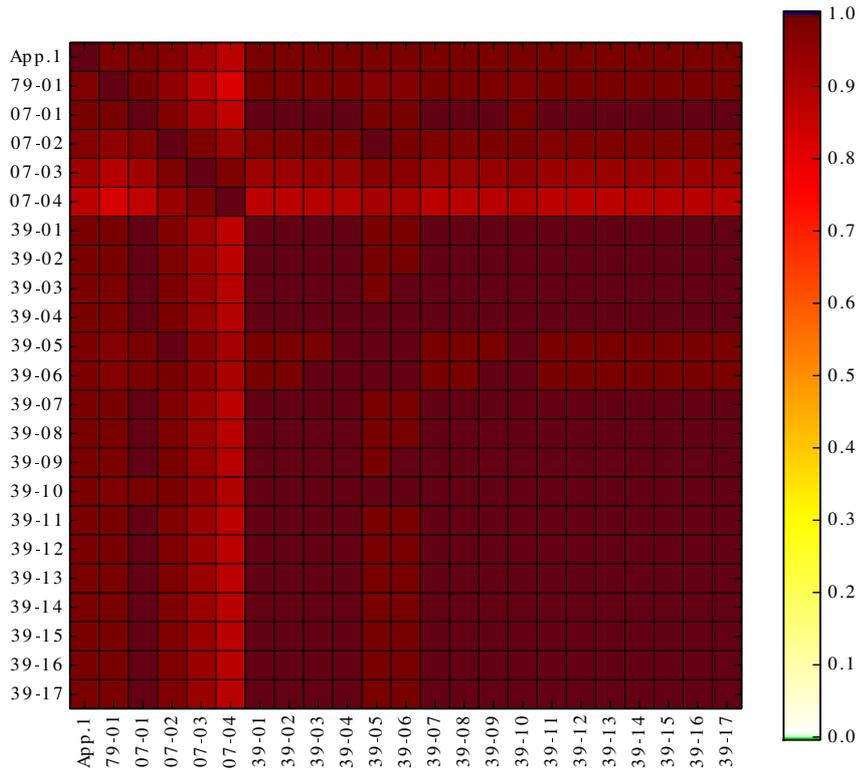
Fig. 4.41   Color coded correlation coefficients due to nuclear data for the application cases App.1 and LCT-79-01 and the experimental data.

COMP-THERM-79-01 taken from the ICSBEP handbook. The configuration consists of hexagonal pitched UO_2 fuel rods, physically similar to the LEU-COMP-THERM-07 and -39 series, but completely independent in terms of considered uncertainties of the system parameters. With exception of LEU-COMP-THERM-07 cases 3 and 4, the k_{eff} values are almost fully correlated. The exception is due to moderation effects as already mentioned in section 3.3. The matrix shown in fig. 4.41 represents the covariance matrix $cov^{0,A}$ in equation 2.22 for the two application cases App1. and LCT-79-01 and the Benchmark experiments.

The correlation coefficients due to system parameters are shown in fig. 4.42 for two different modeling approaches: scenarios A and E, as discussed in section 3.3.1 and were identified as the bounding cases with respect to the resulting correlation coefficients. Scenario A leads to almost full correlation of the data due to system parameters, except for LCT-07 03 and 04. The latter shows significantly lower correlation coefficients. The right
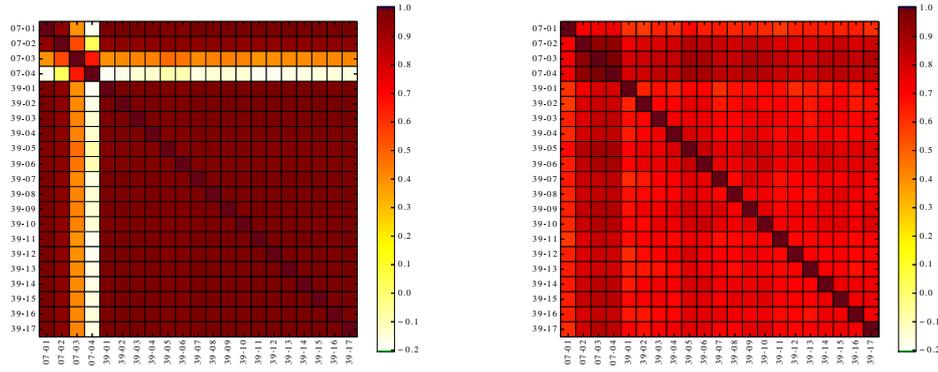
Fig. 4.42   Color coded correlation coefficients for the selected 21 Benchmark experiments for scenario A (left) and E (right).

picture shows the correlation coefficients resulting from scenario E. All coefficients are lower compared to scenario A, except LCT-07 cases 2, 3, and 4, which are almost fully correlated. The differences of the LCT-07 cases 2, 3, and 4 are due to moderation and thus neutron spectrum effects. The described covariance matrices were used in the updating process. The results are shown in fig. 4.43. The figure shows the k_{eff} distributions
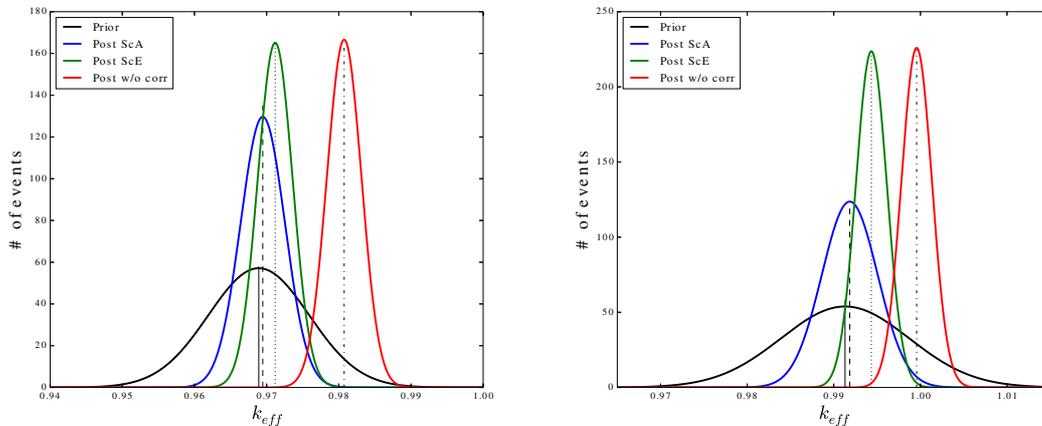
Fig. 4.43   Prior and posterior distributions for the application cases App1 (left) and LCT-79-01 (right). The posterior distributions are shown for scenarios A and E and neglecting systematic uncertainties.

for App.1 and LCT-79-01 for the prior and posterior distributions. The latter distribution is calculated for the two scenarios A and E, and for the case of neglecting the systematic uncertainties (w/o corr). The numerical values are given in table 4.12. The correlation coefficients decrease from scenario A to E to the w/o corr scenario. The effect on the pos-

Tab. 4.12 Resulting k_{eff} values and the 1-σ deviations for prior and different posteriors.

	Prior	Post ScA	Post ScE	Post w/o corr
App.1	0.96891	0.96946	0.97118	0.980743
	$\pm 6.9830 \times 10^{-3}$	$\pm 3.0770 \times 10^{-3}$	$\pm 2.4153 \times 10^{-3}$	$\pm 2.3940 \times 10^{-3}$
LCT79-01	0.9913	0.99184	0.99434	0.99957
	$\pm 7.4045 \times 10^{-3}$	$\pm 3.2262 \times 10^{-3}$	$\pm 1.7835 \times 10^{-3}$	$\pm 1.7656 \times 10^{-3}$

terior distribution can be seen in both pictures in fig. 4.43: the increase in the correlation coefficient pushes the distribution mean further away from the prior mean and decreases the 1-σ deviation. This means for the posterior mean of LCT-79-01 in the right picture to be closer to the experimental value given in the ICSBEP handbook (0.9999$\pm$0.0016). In the case w/o corr this means using more information than the actual experiments offer. In this sense, the resulting posteriors for the application cases underestimate the 1-σ deviation and overestimate the actual k_{eff}. From the viewpoint of criticality safety assessment, overestimating the k_{eff} with a too small 1-σ deviation could be named a conservative approach. The smaller the resulting correlation coefficients were, the larger the shift of the mean value and the smaller the standard deviation becomes. In terms of information used in the updating process, the smaller the correlation coefficients for the system parameters are, the more information can be used to update the prior k_{eff} distribution. In turn, correlation coefficients of almost one means, that the experimental setups are almost identical and the amount of new information in the updating process is tiny.

In a validation process, the desired values of the correlation coefficients between experiments due to manufacturing tolerances are contrary to the correlation coefficients, or c_k values by TSUNAMI, calculated from the variation of nuclear data between the application case and each experimental setup. These correlation coefficients are desired to be close to one to be sure of describing the same nuclear system.

4.2.1.1   Summary and Discussion

We summarize that neglecting the statistical dependence between the system describing parameters of the experimental setups leads to the largest k_{eff} and in light of the before mentioned arguments to an overestimating. The opposite is true for the standard deviation: Neglecting correlation in the data leads to a smaller uncertainty. The analysis

of the data assumes more information than available. However, the SCALE calculations with the used library and the CSAS5 sequence under estimates k_{eff} in general, a known effect for low enriched uranium setups. In this case, the ignoring of correlations leads to an overestimation of k_{eff}, an acceptable circumstance in the light of criticality safety considerations. This might change for codes and setups where the calculated k_{eff} is over estimated compared to experimental values. Neglecting the correlations in the data of a validation processes might lead to an unjustified lowering of the posterior k_{eff}.

It was shown that depending on modeling assumptions, the same data sets lead to different covariance and correlation matrices and in turn to different posterior k_{eff} distributions. In this sense, the resulting difference of the distributions can be viewed as additional uncertainty in the bias prediction.

For more general statements on the impact of correlated data due to system parameters in the process of validation more effort is needed. One crucial point in the analysis is the interpretation of the uncertainty data given in the experimental description.

4.2.2   Example 2: Varying Distributions of Input Parameters

The benchmark experiments and the corresponding covariance and correlation matrices described in section 3.4 are now used for the Bayesian updating process of the application case App.1 described above. To demonstrate the effect of a positive and negative bias on the updating process, the data set is split into two subsets. Subset 1 comprises the five experiments from series LCT-06 and 35 and subset 2 contains the four experiments from series LCT-62. The results are summarized in fig. 4.45. The upper plot shows results for parameters assumed with normal probability distributions, the upper right plot for all parameters with uniform distributions and the lower plot with parameter distributions according to expert judgment. The distribution characteristics of the prior and posterior functions are summarized in table 4.13.

Although the correlation coefficients for the three cases considered here change by up to 20% the effect in bias determination is minor, as can be seen in table 4.13. The respective k_{eff} mean values and the standard deviations are essentially the same for each subset. The Bayesian updating process using the two subsets of benchmark experiments and omitting correlations results in a small shift of the mean and by almost a factor of two
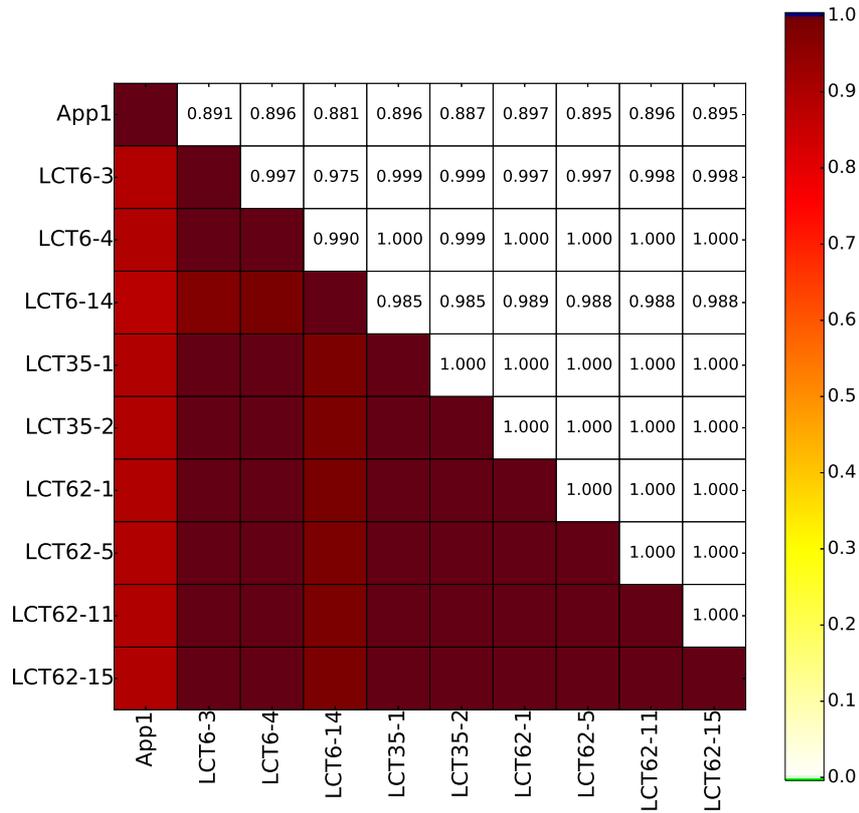
Fig. 4.44 Tsunami c_k values for the application case App1 and benchmark experiments. All values are close to 1, indicating a high similarity.

smaller line width of posterior distributions. As expected, the posterior shifts with subset 1 to larger and with subset 2 to smaller k_{eff} values. Including correlations reduces the amount of the shift so the mean values of the posterior distributions are closer to the mean of the prior. Within the 1-σ range of the posterior distributions their mean values agree with the mean of the prior.

For criticality safety assessment a prediction of the effective neutron multiplication factor including a 95% confidence interval is required to be below a sufficient safety margin. Applying the Bayesian updating process with subsets 1 and 2 results in smaller 95% confidence interval values compared to the prior (see fig. 4.45). The confidence interval value reduces from 0.980 of the prior to 0.976 of the posterior with subset 1 and to 0.974 with subset 2 and correlations included. However, in this combination of benchmark experiments and application case the prior represents the conservative approach.
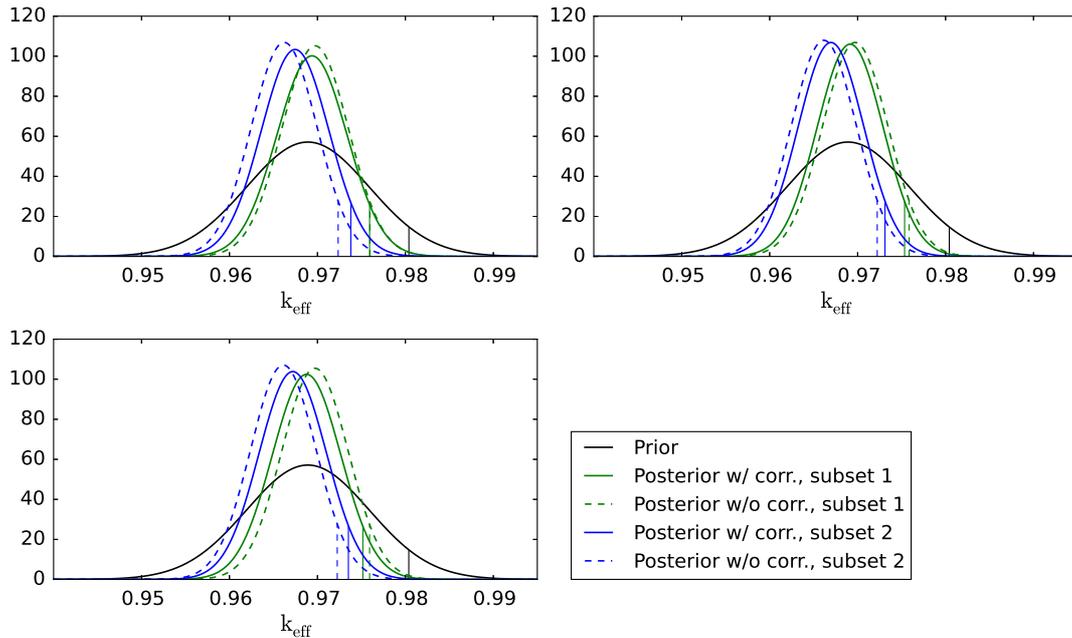
Fig. 4.45   Prior and posterior distributions for the application case from [11] using experiments from LCT-06 and 35 in subset 1 and from LCT-62 in subset 2 as benchmark experiments and assuming that all parameter distributions are normal (upper left plot), uniform (upper right plot) and a mixture of both (lower plot). The 95% safety limits are indicated by vertical lines.

Now the Bayesian updating process is applied assuming experiment LCT-62-15 as the application case. For the updating process, experiments from series LCT-07 and 39 and the corresponding covariance matrix are used as described in section 3.3. The c_k-values as predicted by TSUNAMI are shown in figure 4.46. The prior and the resulting posterior distributions with and without correlations are shown in fig. 4.47. Here, the calculated prior of the effective neutron multiplication factor $k_{\mathrm{eff}}^{\mathrm{prior}} = 1.00353 \pm 0.00642$ of the applications case is overestimated by SCALE, while the k_{eff} values of the benchmark experiments are all underestimated. In this combination the posterior mean value shifts by the updating process to $k_{\mathrm{eff}}^{\mathrm{post}} = 1.00481 \pm 0.00263$ with correlations and to $k_{\mathrm{eff}}^{\mathrm{post}} = 1.00596 \pm 0.00178$ without correlations and thus even further away from the experimental value $k_{\mathrm{eff}}^{\mathrm{exp}} = 1$. Simultaneously the posterior distributions narrow by about a factor of three. Also in this combination of application case and benchmark experiments the k_{eff} mean values of the posterior agree within their standard deviation with the mean value of the prior. The 95% confidence interval value, however, decreases from 1.0141 of the prior to 1.0091 of the posterior with correlations and to 1.0089 without correlations. Also in this example the prior represents the conservative approach.

Tab. 4.13 Posterior distribution characteristics ($k_{\text{eff}} \pm 1\sigma$) for the prior N(0.96891, 0.00698) assuming correlation (w/ cor) and neglect them (w/o cor) and for the subsets 1 and 2.

	normal distributed	uniform distributed	expert judgment
post w/ cor, subset 1	0.96938 ± 0.00398	0.96915 ± 0.00376	0.96876 ± 0.00389
post w/o cor, subset 1	0.96973 ± 0.00379	0.96973 ± 0.00376	0.96972 ± 0.003478
post w/ cor, subset 2	0.96746 ± 0.00373	0.96697 ± 0.00373	0.96720 ± 0.00384
post w/o cor, subset 2	0.96621 ± 0.00373	0.96615 ± 0.00370	0.96612 ± 0.00373

Using subsets of experiments with under- and overestimated k_{eff} respectively leads to different behavior in the Bayesian updating process. Updating of an underestimated application case with overestimated benchmark experiments or vice versa shifts the mean of the posterior distribution even further away from the experimental value, as expected. In case of an overestimated application case the posterior will shift to more conservative values. However, within one standard deviation the mean values of the posterior distributions overlap with the mean value of the prior. The updating process not only shifts the mean value but also narrows the posterior distribution. Determining the 95% confidence interval results in smaller values for the posterior distributions compared to the prior. Criticality safety assessment requires a prediction of k_{eff} below a safety margin. Applying the updating process, with experiments described here, the 95% confidence interval of the posterior distributions may fulfill this criterion while the prior distribution would fail. However, the prior represents the more conservative approach.

4.2.2.1   Summary and Discussion

Varying the distribution functions of system parameters also alters the correlation coefficients between the calculated k_{eff}. Although some of the correlation coefficients change by up to 20% for the three cases and benchmark experiments considered here, their impact on bias estimation is minor. This demonstrates that for these experiments the choice of distribution functions of system parameters is not crucial for bias estimation. Nevertheless, the choice of the distribution functions should be chosen and justified based on expert judgment. Furthermore, the results demonstrate that benchmark experiments as well as the application case need to be chosen and analyzed carefully. Some benchmark experiments might even be discarded in further analysis, if they do not reflect the appli-
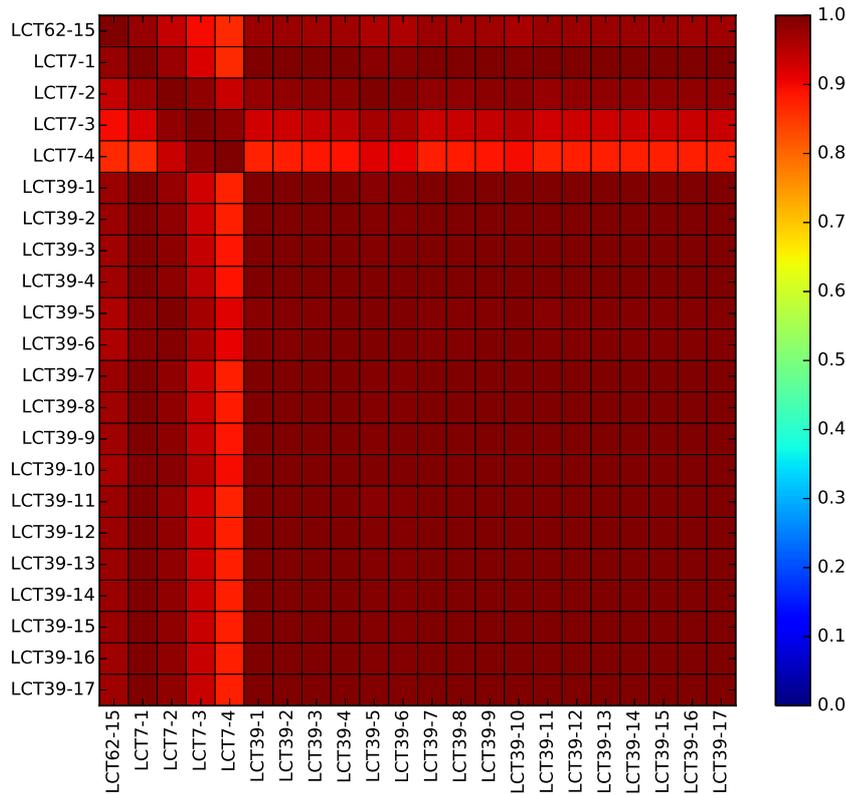
Fig. 4.46 Tsunami c_k values for the application case LCT62-15 and benchmark experiments from LCT07 and -39. All values are close to 1, indicating a high similarity.

cation case in a sufficient way. A similarity of the neutron spectrum properties between the experiments and the application case is only one criterion. Another may be e.g. geometrical similarities, or whether the same type of fuel rods is used. Hidden experimental and/or calculation systematic error in the benchmark experiments as well as in the application case may lead to misleading results. However, shifts to larger effective neutron multiplication factors represent conservative approaches.

Again, all results and discussions presented here only apply to LEU-COMP-THERM experiments calculated with SCALE6.1.2 in combination with ENDF/B-VII continuous energy library. For more general and detailed statements further investigations are required.
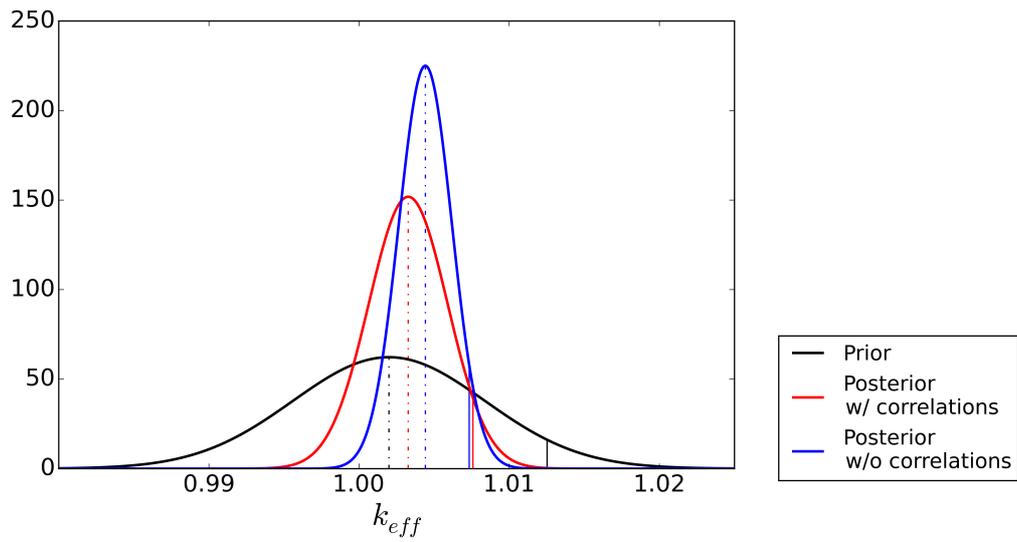
Fig. 4.47 Prior and posterior distributions for LCT-62-15 as application case and experiments from the series LCT-07 and -39 described in section 3.3 as benchmark experiments. The vertical dash-dotted lines indicate the mean values and the solid lines the 95% confidence intervals.

5   Concluding Remarks

This work dealt with the question of how to treat statistical dependencies within data of experimental series of critical experiments if this data is used for validating purposes. Statistical dependencies or correlation effects can arise if parameters, materials or parts of the experimental setups are shared within experimental series.

The work seems necessary due to the increasing computer power which in combination with the availability of user friendly sophisticated software and the publicly available huge collections of evaluated experimental data can be used to perform massive analysis. For a validation this could lead to the use of not just 10 or 20 experiments, but hundreds without great effort. This would need a correct translation of the available data into numerical models taking into account the correct statistical treatments. The latter needs the consideration of integral experimental covariance data, which has to be generated. However, even the consideration of only few but correlated experiments makes the consideration of integral experiment covariances necessary, too.

We have used a method based on full Monte-Carlo sampling to investigate the generation of covariance and the related correlation matrices (section 3). The Monte-Carlo sampling method provides some advantages in terms of being easy to understand and taking nonlinear dependencies into account. The major drawback is the enormous computational effort. In section 3.3 we have shown, that a sample size up to 200 is needed to achieve convergence, leading to the need of a computer cluster to calculate covariance matrices for a few experiments. More sophisticated Monte-Carlo approaches which need less samples (e.g. the Latin-Hypercube Monte-Carlo sampling) might reduce the calculation times significantly.

However, having enough time and computer power, this problem can be solved. More difficult to circumvent is the obstacle of translating the actual experimental data as given for example in [1] into calculation models. The experimental data is not always complete regarding the statistical dependence of parameters and thus leaves some freedom in the interpretation of statistical dependencies and modeling. This leads in turn to different uncertainties of k_{eff} (e.g. fig. 3.10) and different sensitivities of the uncertainty on the input parameters (e.g. fig. 3.16) and thus to different covariance matrices (e.g. fig. 3.13). To generate reliable covariance matrices, the treatment of data gaps has to be clarified.

That might lead to some exclusions of data sets. Even though a lot of data is available and publicly accessible, not all data might be used for the generation of reliable covariance data due to lacking information.

For the consideration of the covariance data within a validation procedure we described and investigated a method based on the Bayesian updating process (section 4). We have shown, that depending on the application case and the used experimental data, ignoring correlation effects can have a significant impact on the validation. However, so far we would not be able to give general statements of how to treat integral experimental covariance data in the process of validation. All our conclusions drawn are only valid for the investigated cases of application case and benchmark experiments.

The final conclusion for this work are:

– If statistical dependencies can be excluded (e.g., selection of experiments each from different series conducted in different laboratories), all presented issues can be circumvented except for the modeling and interpretation of the experimental data.

– If statistical dependencies exist in the experimental data used for validation, the integral experimental covariances have to be taken into account. If the data allows for different covariance matrices, the most conservative case should be considered. This depends on the application, the benchmark experiments and the calculation assumptions used (modeling, nuclear data, solution methods of the transport equation, etc.).

– The presence of shared components within experimental series does not necessarily lead to statistically significant correlation coefficients. If the shared experimental components have comparable small uncertainties or play only a minor role for the k_{eff} uncertainty (determined e.g. by means of a sensitivity analysis), their contribution to the correlation coefficient is negligible.

– Modeling of experimental data for the calculation of integral covariance data must be investigated and described further.

– The handling of data gaps in the description of used experimental data has to be investigated and described further.

References

[1] "International handbook of evaluated criticality safety benchmark experiments," Tech. Rep. NEA/NSC/DOC(95)03, NEA Nuclear Science Committee, 2015.

[2] T. T. Ivanova, M. N. Nikolaev, K. F. Raskach, E. V. Rozhikhin, and A. M. Tsiboulia, "Influence of the correlations of experimental uncertainties on criticality prediction," *Nuclear Science and Engineering*, vol. 145, 2003.

[3] M. Bock and M. Stuke, "Determination of correlations among benchmark experiments by monte carlo sampling techniques," *Proc. ANS Nuclear Criticality Safety Division Topical Meeting (NCSD 2013), Wilmington, North Carolina, USA*, 2013.

[4] M. Bock and M. Behler, "Impact of correlated benchmark experiments on the computational bias in criticality safety assessment," *Proc. ANS Nuclear Criticality Safety Division Topical Meeting (NCSD 2013), Wilmington, North Carolina, USA*, 2013.

[5] E. Peters, F. Sommer, and M. Stuke, "Sensitivity and correlation of critical experiments due to manufacturing tolerances and cross section uncertainties," *Proc. ICNC 2015, Sept. 13 - 17, 2015, Charlotte, North Carolina, USA*, 2015.

[6] E. Peters, F. Sommer, and M. Stuke, "Impact of correlated data in validation procedures," *Proc. ICNC 2015, Sept. 13 - 17, 2015, Charlotte, North Carolina, USA*, 2015.

[7] A. Hoefer, O. Buss, M. Hennebach, M. Schmid, and D. Porsch, "MOCABA: a general Monte Carlo-Bayes procedure for improved predictions of integral functions of nuclear data," *Annals of Nuclear Energy*, 2014. 10.1016/j.anucene.2014.11.038.

[8] V. Sobes, B. T. Rearden, D. E. Mueller, W. J. Marshall, J. M. Scaglione, and M. E. Dunn, "Upper subcritical calculations based on correlated data," *Proc. ICNC 2015, Sept. 13 - 17, 2015, Charlotte, North Carolina, USA*, 2015.

[9] C. Baker, P. N. Smith, R. Mason, M. Shepherd, S. Richards, R. Hiles, R. P. adn D. Hanlon, and G. Dobson, "Calculating uncertainty on k-effective with monk10," *Proc. ICNC 2015, Sept. 13 - 17, 2015, Charlotte, North Carolina, USA*, 2015.

[10] W. J. Marshall and B. T. Rearden, "Determination of critical experiment correlations using the sampler sequence within scale 6.2," *Proc. ICNC 2015, Sept. 13 - 17, 2015, Charlotte, North Carolina, USA*, 2015.

[11] A. Hoefer, O. Buss, T. Ivanova, B. Rearden, W. Marshall, D. Mennerdahl, and M. Stuke, "Proposal for benchmark phase iv: Role of integral experiment covariance data for criticality safety validation," *OECD-NEA WPNCS EGUACSA*, 2016.

[12] K. Olive *et al.*, "2015 review of particle physics," *Chin. Phys. C*, vol. 38, 2014 and 2015 update.

[13] C. A. Peters, "Statistics for analysis of experimental data," in *Environmental Engineering Processes Laboratory Manual* (S. E. Powers, ed.), Champaign, IL: AEESP, 2001.

[14] R. Fisher, "Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population," *Biometrika (Biometrika Trust)*, vol. 10 (4), 1915.

[15] J.Neuber and A.Hoefer, "Frequentist and bayesian approach in criticality safety uncertainty evaluations," *Proc.2009 Nuclear Criticality Safety Division Topical Meeting NCSD 2009, September 13-17, 2009, Richland, Washington(USA)*, 2009.

[16] A. Hoefer, O. Buss, and J. Neuber, "How confident can we be in confidence intervals for the computational bias obtained with the generalized linear least squares methodology?," *Proc. International Conference on Nuclear Criticality (ICNC 2011). Edinburgh, Scotland*, 2011.

[17] "Scale: A comprehensive modeling and simulation suite for nuclear safety analysis and design," Tech. Rep. ORNL/TM-2005/39, Version 6.1.2, available from RSICC at Oak Ridge National Laboratory as CCC-785, Oak Ridge National Laboratory, 2012.

[18] O. Buss, A. Hoefer, and J. Neuber, "Hierarchical monte carlo approach to bias estimation for criticality safety calculations," *Proc. Physics of Reactors International Conference (PHYSOR 2010), Pittsburgh, Pennsilvania, USA*, 2010.

[19] M. Williams and B. Rearden, "Scale-6 sensitivity/uncertainty methods and covariance data," *Nuclear Data Sheets*, vol. 109, no. 12, pp. 2796 – 2800, 2008. Special Issue on Workshop on Neutron Cross Section Covariances June 24-28, 2008, Port Jefferson, New York, {USA}.

[20] C. J. Diez, O. Buss, A. Hoefer, D. Porsch, and O. Cabellos, "Comparison of nuclear data uncertainty propagation methodologies for PWR burn-up simulations," *Annals of Nuclear Energy*, 2014. 10.1016/j.anucene.2014.10.022.

[21] W. Zwermann, L. Gallner, M. Klein, B. Krzykacz-Hausmann, I. Pasichnyk, A. Pautz, and K. Velkov, "Status of xsusa for sampling based nuclear data uncertainty and sensitivity analysis," *EPJ Web of Conferences*, vol. 42, 2013.

[22] M. Behler, M. Bock, F. Rowold, and M. Stuke, "Suncistt - a generic code interface for uncertainty and sensitivity analysis," *Proc. Probabilistic Safety Assessment and Management PSAM12, Honolulu, Hawaii, USA, 22-27 June*, 2014.

[23] M.Kloos, "Susa version 4.0, software for uncertainty and sensitivity analyses, user's guide and tutorial," Tech. Rep. GRS-P-5, Rev. 2, Gesellschaft fuer Anlagen- und Reaktorsicherheit (GRS) mbH, Garching, Germany, 2016.

[24] E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," 2001-2016. [Online; accessed 2016-10-01].

[25] "Root - an object oriented data analysis framework," *Phys. Res. A 389 (1997) 81-86*, 1997.

[26] W. J. Marshall and B. T. Rearden, "Criticality safety validation of scale 6.1," Tech. Rep. ORNL/TM-2011/4502011, Oak Ridge National Laboratory, 2011.

[27] J. Bonnet, D. Doutriaux, P. Grivot, and G. Poullot, "Programme 1975/1994, crayons u(4.738)o2 gainés ags 'compléments d'informations'," Tech. Rep. IPSN/SRSC/98.03 and IPSN/SEC/T/ 98.424.- Revision B du 22/03/2002, Laboratoire de criticité de Valduc, 2002.

[28] J. Osborne and A. B. Costello, "Sample size and subject to item ratio in principal components analysis," *Practical Assessment, Research & Evaluation*, vol. 9(11), 2004.

[29] L. Aleamoni, "The relation of sample size to the number of variables in using factor analysis techniques," *Educational and Psychological Measurement*, vol. 36, 1976.

[30] E. Guadagnoli and W.F.Velicer, "Relation of sample size to the stability of component patterns," *Psychological Bulletin*, vol. 103, 1988.

[31]   R. Kilger, F. Sommer, and M. Stuke, "Analysis of experimental series of plutonium nitrate in aqueous solution and their correlation coefficients," *Annals of Nuclear Energy*, 2016. 10.1016/j.anucene.2016.06.017.

[32]   E. Peters, F. Sommer, and M. Stuke, "Modeling of critical experiments and its impact on integral covariance matrices and correlation coefficients," *Annals of Nuclear Energy*, 2016. 10.1016/j.anucene.2016.02.011.

[33]   "Database for the international criticality safety benchmark evaluation project," Tech. Rep. September 2015 Edition, build 2.7, OECD-NEA, 2015. DICE.

List of Figures

85

List of Tables