Nuclear Engineering and Technology 54 (2022) 790–802

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

journal homepage: www.elsevier.com/locate/net

Original article

A methodology for uncertainty quantification and sensitivity analysis for responses subject to Monte Carlo uncertainty with application to fuel plate characteristics in the ATRC

Dean Price*, Andrew Maile, Joshua Peterson-Droogh, Derreck Blight

Idaho National Laboratory, P.O. Box 1625, Idaho Falls, ID, 83415-3885, USA

ARTICLE INFO

Article history:
Received 23 June 2021
Received in revised form
10 August 2021
Accepted 8 September 2021
Available online 11 September 2021

Keywords:
Advanced test reactor critical
Advanced test reactor
ATRC
Uncertainty quantification
Criticality
Sensitivity analysis
Monte Carlo

ABSTRACT

Large-scale reactor simulation often requires the use of Monte Carlo calculation techniques to estimate important reactor parameters. One drawback of these Monte Carlo calculation techniques is they inevitably result in some uncertainty in calculated quantities. The present study includes parametric uncertainty quantification (UQ) and sensitivity analysis (SA) on the Advanced Test Reactor Critical (ATRC) facility housed at Idaho National Laboratory (INL) and addresses some complications due to Monte Carlo uncertainty when performing these analyses. This approach for UQ/SA includes consideration of Monte Carlo code uncertainty in computed sensitivities, consideration of uncertainty from directly measured parameters and a comparison of results obtained from brute-force Monte Carlo UQ versus UQ obtained from a surrogate model. These methodologies are applied to the uncertainty and sensitivity of k_{eff} for two sets of uncertain parameters involving fuel plate geometry and fuel plate composition.

Results indicate that the less computationally-expensive method for uncertainty quantification involving a linear surrogate model provides accurate estimations for k_{eff} uncertainty and the Monte Carlo uncertainty in calculated k_{eff} values can have a large effect on computed linear model parameters for parameters with low influence on k_{eff}.

© 2021 Korean Nuclear Society, Published by Elsevier Korea LLC. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

In practical engineering scenarios, particularly those involving experimental implementation, it is important to have an in-depth understanding of the behavior and consistency of a computational model under perturbations to model parameters that lie within a realistic bound of tolerances. This allows engineers to gain more insight into the validity of the model when comparing results to experimental results. One way of gaining this understanding is through uncertainty quantification (UQ) and sensitivity analysis (SA). Uncertainty quantification can be broadly defined as the "science of identifying, quantifying and reducing uncertainty associated with models" Smith [1] and sensitivity analysis can be described as the "study of how uncertainty in the output of a model can be apportioned to different sources of uncertainty in model input" Saltelli et al. [2]. Often, the definition of uncertainty analysis can be said to end before taking action to reduce those uncertainties. Nevertheless, when combined, these two routes of exploration offer a comprehensive understanding of the relationship between model input and output as well as a confidence bound on the results of the model.

Overall, sources of uncertainty can affect physical models at every step in the calculation. A reasonable classification system for uncertainty sources is presented in Radaideh and Kozlowski [3]. This paper describes parametric uncertainty as the result of "stochastic or unknown" behavior in model parameters. A variety of methods exist to propagate parametric uncertainty through a physical model in order to characterize the behavior of an output parameter. One popular method involves using Monte Carlo sampling techniques [4]. performed a study to test various sampling strategies in the context of Monte Carlo uncertainty analysis. The model used for analysis had a civil engineering application involving the calculation of heat and moisture transfer between building components; it was found that the more sophisticated sampling strategies increased the efficiency of Monte Carlo based uncertainty analysis. In the field of nuclear engineering, techniques such as "Fast Total Monte Carlo" or "Fast Generation Random Sampled" have been developed to propagate uncertainties through

* Corresponding author.
E-mail address: dean.price@inl.gov (D. Price).

https://doi.org/10.1016/j.net.2021.09.010
1738-5733/© 2021 Korean Nuclear Society, Published by Elsevier Korea LLC. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Monte Carlo calculations with small additional compute time Rochman et al. [5]. Deterministic methods also exist for uncertainty propagation by solving adjoint equations for the underlying physics of a problem; a popular tool for this in the nuclear community is the TSUNAMI-3D code Rearden et al. [6] which can perform adjoint-based eigenvalue sensitivity analysis for cross-sectional data. Another study Price et al. [7] used both deterministic methods with TSUNAMI-3D and Monte Carlo methods to evaluate the effect of nuclear data uncertainty on single lattice k_{inf}.

A useful overview for sampling based UQ/SA methods and how these methods typically align to generate a complete workflow can be found in this survey study Helton et al. [8]. One study that implements a few different SA methods with application to the performance of solid oxide fuel cells as a sample problem for comparison of methods can be found at Radaideh and Radaideh [9]. Results from direct perturbation (called “one at a time” in this study), Morris screening, standardized regression coefficients and partial correlation coefficient were compared. Nonlinear methods such as standardized rank regression coefficients and partial rank correlation coefficients were used as well.

The present study includes parametric UQ/SA on the Advanced Test Reactor Critical (ATRC) facility housed at Idaho National Laboratory (INL). Due to the availability of data and demand for directly-applicable results, this study offers a unique approach to UQ/SA which includes consideration of Monte Carlo code uncertainty in computed sensitivities, consideration of uncertainty from directly measured parameters, a comparison of results obtained from brute-force Monte Carlo UQ versus UQ obtained from a surrogate model and finally both linear and nonlinear methods for SA. Furthermore, parametric SA is often difficult to perform using Monte Carlo based physics simulations due to the statistical uncertainty in the final result. Overall, this can make parametric SA difficult to perform using complex methods on high-fidelity models. Due to the large amount of computational resources dedicated to this study, SA is performed and a method for quantifying the effect of statistical uncertainty in the output on computed parametric sensitivities is presented. The effect of uncertainty in fuel plate composition and fuel plate geometric parameters on reactor k_{eff} are considered. The uncertainty in these parameters is obtained directly from data reported by direct measurements or fuel vendors, another unique aspect of this study is that sampled parameters may not be those explicitly represented in the neutronics model (i.e. sampling B_4C mass in fuel and B-10 abundance as separate uncertain parameters leads to an implicit distribution of B-10 included in the sampled models). The methodology described in this paper can act to guide UQ and SA efforts on expensive computational models that use Monte Carlo methods for calculation.

2. Data resources

This section presents information on the resources used to carry out this study. First, a description of the full ATRC base model (made for the neutronics code MC21) is given with a focus on characteristics pertaining to the fuel elements. Next, information on the computer code and cross-section data are given. Finally, uncertainty sources and distributions for the two groups of uncertain parameters are presented: (1) fuel geometry uncertainty and (2) fuel composition uncertainty.

2.1. Description of the base model

The ATRC reactor is an open-pool serpentine reactor that uses highly-enriched uranium fuel plates. The purpose of the ATRC is to test core design experiments in terms of their reactivity characteristics and axial flux profiles prior to their irradiation in the Advanced Test Reactor (ATR). A unique feature of the ATRC is that reactivity is controlled with beryllium rotating cylinders (called ”outer shim control cylinders”). This allows for the axial flux profile of the reactor to be minimally affected by the reactivity control mechanisms. The model used in this analysis was created to reduce the cost associated with ATRC operation by providing estimates on important quantities in various experimental configurations to support experimentally-observed results; it is made for the MC21 neutronics code Sutton et al. [10]. Due to the desired accuracy of the results, the model is a 3-D full core model. This likely has a minimal effect on the results because of the high axial uniformity present in the ATRC core design. To provide some more information, all irradiation positions for experiments span the axial height of the core. As mentioned previously, the control cylinders also span the axial height of the core. The only control element which does not span the height of the core are the neck shim rods. The other exception to the high degree of axial homogeneity present in the real ATRC are the experiments that are placed in the irradiation positions. Often, these experiments are not uniform in the axial direction. However, in the model used in this analysis, no experiments are modeled in the irradiation positions so this effect is not present. Also, no symmetry such as quarter-core or half-core symmetry is assumed—despite that symmetry existing in the model.

Fig. 1 is included to give the reader a general idea of the design of the ATRC reactor. In this diagram, there are number of experimental irradiation positions for samples to be placed. As mentioned earlier, the ATRC functions to quantify the reactivity insertion of experimental samples—not to provide the high-flux environment to actually irradiate these samples. Therefore, a unique opportunity exists to perform the present analysis on fuel with negligible burnup. It is not necessary to rely on computational predictions of fuel isotopics which may incorporate systematic uncertainties that may bias the results of the present work. The fuel elements are arranged in a “serpentine” pattern in the core, they wrap around the central experimental irradiation positions. Each element consists of 19 curved fuel plates composed of an aluminum outer cladding and inner fuel meat. Finally, two of three means of reactivity control are shown in this diagram, the outer shims and the

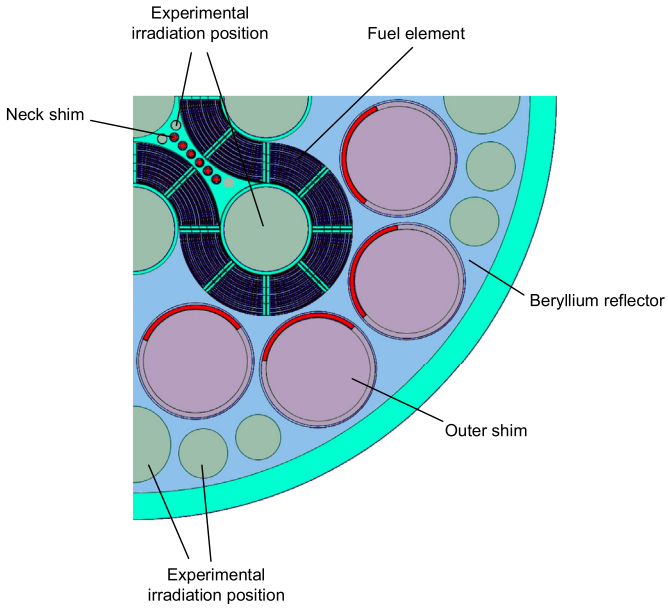

Fig. 1. General radial layout ATRC core.

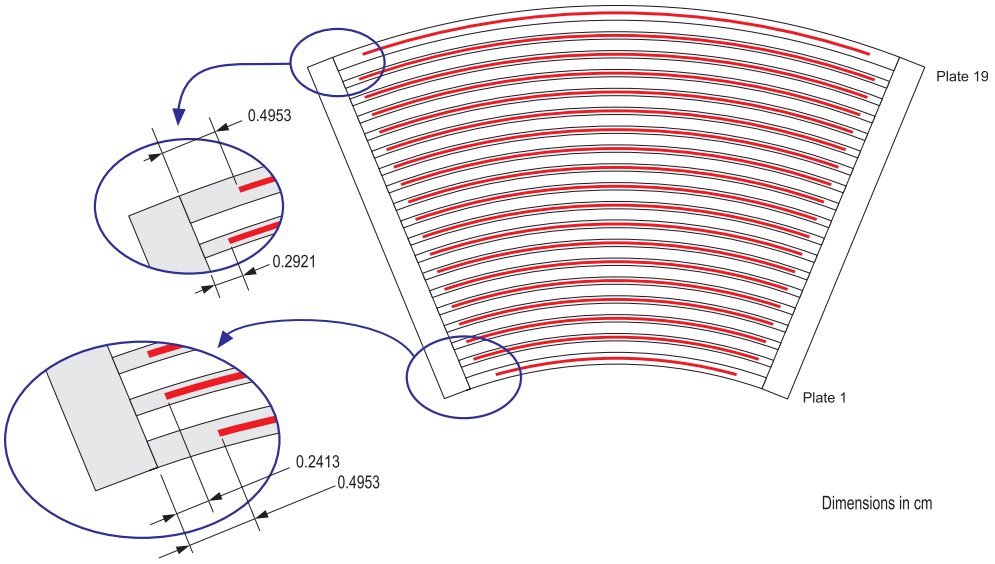

Fig. 2. Diagram of fuel elements in ATRC, adapted from Kim and Schnitzler [12].

Table 1
Distributions for uncertain fuel geometry parameters.

Parameter	Mean [cm]	Standard Deviation [cm]	95% CI [cm]
Inset distance of plate 1	0.4953	0.0324	±0.0635
Inset distance of plates 2-17	0.2413	0.0324	±0.0635
Inset distance of plate 18	0.2921	0.0324	±0.0635
Inset distance of plate 19	0.4953	0.0324	±0.0635
Fuel meat height of plates 1-19	121.92	0.9719	±1.905

neck shims. The outer shims are the primary method of reactivity control and rotate to either expose hafnium neutron poison or beryllium reflector to the body of the core. The reactivity control mechanisms not shown in the figure are the safety rods.

For this analysis, a few assumptions are made for this model. First, for models used in the part of this study where uncertainties in fuel composition are concerned, the fuel composition of the 38 perturbed fuel elements is homogenized. Due to the detail of measured data available for the design of 38 out of 40 of the fuel designs, the compositions of only 38 out of 40 fuel elements are perturbed in the models used for the fuel composition study. Next, the base loading configuration was used, it is thought that the critical configuration of fuel elements and reactivity controls will have a small effect on the observed results. Due to the insignificant power output of the reactor, all material temperatures are kept at 300K.

2.2. Computational tools

As mentioned in Section 2.1 the MC21 Monte Carlo transport code is used in Sutton et al. [10]. Capable of both neutron and photon transport, the code is under continuing development at the Naval Nuclear Laboratory. Continuous neutron energy cross-sections are used in this study calculated using the ENDF/B-VII.1 nuclear data libraries Chadwick et al. [11].

For the first group of uncertain parameters, the fuel geometric parameters, the model was run with a minimum of 1500 batches, 50 discard and 300,000 histories. Due to the large model size and computational demand of the study, these Monte Carlo parameters had to be adjusted according to the availability of computational resources. Overall, the average Monte Carlo standard deviation arising in k_{eff} for the models run with perturbed fuel geometric parameters was 3.3 pcm with a maximum of 4.0 pcm. For the second group of uncertain parameters, the fuel composition parameters, the model was run with a minimum of 3000 batches, 50 discard and 800,000 histories. Overall, the average Monte Carlo

Table 2
Distributions for uncertain fuel geometry parameters.

Parameter	Mean	Standard Deviation	95% CI
Uranium-aluminide mass in element [g]	1467	9.86	19.3
Boron-carbide mass in element [g]	5.280	0.054	0.106
Aluminum (X8001) mass in element [g]	1288	15.7	30.8
Total uranium mass in element [g]	1045	0.772	1.51
Total uranium-235 mass in element	973.5	0.369	0.723
Boron-10 abundance	0.1965	0.0038	7.50×10^{-3}
Uranium-234 fraction	0.0098	1.00×10^{-4}	1.96×10^{-4}
Uranium-236 fraction	0.0045	4.59×10^{-5}	9.00×10^{-5}

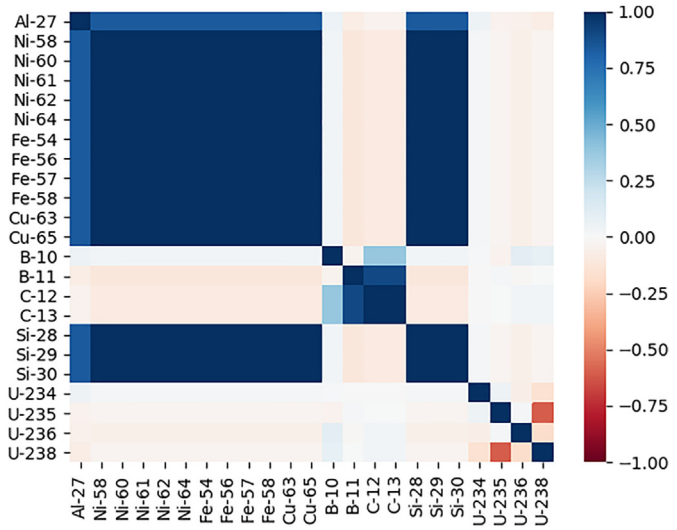

Fig. 3. Implicit correlation between isotopic concentrations from converting sampled quantities into code-input fuel compositions.

standard deviation arising in k_{eff} for the models run with perturbed fuel composition parameters was 1.6 pcm with a maximum of 1.7 pcm. Although these code-reported uncertainties seem small, they can have large impacts on computed sensitivity measures. This problem can lead to difficulty in using stochastic methods for neutron transport in sensitivity analysis, however, this paper presents a method to quantify the impact of these uncertainties on computed sensitivities as detailed in Section 3.3. The computational workflow used to carry out this study is given in Appendix A.. For reference, each full model evaluation took around 1500 core-hrs. All calculations were run on the high performance computing clusters at Idaho National Laboratory. In order to generate the text files provided to the physics code that represent the models, the popular scripting tool Python was used. Python was also used to parse and analyze the output coming from the numerous calculations provided by MC21.

2.3. Fuel geometry uncertainty

Of the two sets used in this study, the first set of uncertain parameters involve aspects of the fuel meat geometry. Fig. 2 depicts a cross-cut of the fuel elements used in the ATRC, with the relevant geometry parameters marked. In this figure, the fuel meat is the red region located at the center of each of the 19 fuel plates. The smallest plate at the bottom of this figure is plate 1 with the largest plate at the top being plate 19. The two parameters within the fuel geometry set that will be varied in this analysis, *across each of the 19 fuel plates independently,* is the fuel inset distance and the fuel meat height. Therefore, a total of 38 independent uncertain parameters are included in this analysis. The fuel elements across the core are kept identical by perturbing each of their 19 plates in the same manner. The inset distance is the distance between the outer edge of the fuel meat and the inner edge of the side plate. The inset distance is measured to be 0.4953 cm for both the smallest and largest fuel plates which lie on the outsides of the fuel element. Plate 18 has a mean inset distance of 0.2921 cm. For the remaining 16 fuel plates this quantity is measured to be 0.2413 cm. The fuel meat height is the total z-height of the fuel meat in the reactor. It is reported to be 121.92 cm across all fuel plates. These parameters were selected for analysis because they directly effect the characteristics of fissile material in the core. This, combined with the high dimensionality of 38 parameters, will make them useful for demonstrating the methods described in this paper.

Uncertainties in the fuel geometry are obtained directly from facility documentation on the fuel elements. Throughout this study, uncertain parameters are assumed to follow a Gaussian distribution with a 95% confidence interval (CI) corresponding to the reported uncertainty. As such, Table 1 gives the mean values and standard deviations of the fuel geometry parameters.

2.4. Fuel composition uncertainty

The second set of uncertain parameters used in this study involve fuel meat composition characteristics. Often, in fuel composition uncertainty analyses, input uncertainty is directly sampled for the isotopic composition of the fuel Radaideh et al. [8,13]. In this study, mostly directly measured parameters, such as mass of uranium and uranium-aluminide mass are independently sampled. This allows for uncertainties to be used that are the direct result of experimental measurements. Also, only 38 of the 40 total fuel elements have a common design and are perturbed with this group of parameters. The fuel meat consists of highly-enriched uranium-aluminide powder mixed into an aluminum alloy (X8001). Boron-carbide is included in the fuel mixture. In total, 5 quantities with reported experimental uncertainty and nominal values are considered as uncertain parameters:

- Uranium-aluminide mass: mass of uranium aluminide powder in fuel element
- Boron-carbide mass: mass of boron carbide in fuel element
- Aluminum (X8001) mass: mass of aluminum X8001 alloy in fuel element
- Total uranium mass: total mass of all uranium isotopes in fuel element
- U-235 mass: mass of U-235 isotope in fuel element

On top of these measured quantities, an additional 3 parameters are considered to have uncertainty based on existing literature:

- Natural boron-10 abundance: this quantity plays an important role in the neutronic behavior of any reactor due to its high absorption cross-sections. It is considered here because there is a moderately large uncertainty in the abundance of this isotope given an arbitrary sample. The uncertainty and assumed mean value for this quantity can be found at Meija et al. [14].
- U-234 fraction: particularly with high-enriched uranium fuel (as is the case in the ATRC), uncertainties in the concentration of this isotope may contribute to some final uncertainty in k_{eff}. Nominal values for this quantity are obtained from McConn et al. [15]; a 2% relative uncertainty is assumed.
- U-236 fraction: same justification as previous entry. Nominal values for this quantity are obtained from McConn et al. [15]; a 2% relative uncertainty is assumed.

Nominal values and distribution parameters are reported in Table 2.

For any model with these sampled parameters, calculations are performed to convert the sampled quantities into an isotopic composition. Following that the concentration of a particular isotope may be dependent on multiple sampled quantities, an implicit correlation can be quantified between isotopes in these models made from sampled parameters. Fig. 3 shows the correlation matrix between isotopes.

3. Methodology

This section details the methods used for SA and UQ. The first two subsections will compare two different methods of uncertainty quantification: the Monte Carlo method and algebraic propagation of uncertainty through a linear model. The advantages of both methods will be discussed as well as the situations where each method is best put to use. Next, details on how to create the linear model are given, as well as a method to quantify the effect of code-reported Monte Carlo uncertainty on linear model parameters. Finally, four methods will be presented for quantifying the contribution that the uncertainty of each input parameter has on the final uncertainty of k_{eff}. Results derived from these methods can guide the analyst in improving their computational models by identifying inputs that are major contributors of response uncertainty.

3.1. Sampling-based UQ using total Monte Carlo

The primary method of in-depth uncertainty quantification for highly complex models is the total Monte Carlo method Helton et al. [8]. To explain this method, first it is necessary to define a model of interest f which takes a number of inputs. These inputs can be written as a vector, $\boldsymbol{x}$, and the model can yield scalar output y. This relationship, for the general model f, is written in mathematical form in Equation (1).

$$f(\boldsymbol{x}) = y \tag{1}$$

In practice, f may be a highly complex model, that may be the result of a very large number of inputs. However, only input parameters whose uncertainty should be propagated into the final result of y are included in the input vector $\boldsymbol{x}$ for this formulation. Next, it is necessary to describe the uncertainty of each parameter in $\boldsymbol{x}$ using some probability distribution, this is shown in Equation (2) where $\mathcal{H}$ is an arbitrary multivariate distribution. In this study, $\mathcal{H}$ is a multivariate normal distribution with means and diagonal covariance matrix described by Tables 1 and 2. Characterization of the uncertainty in $\boldsymbol{x}$ is an important step that should be informed by experimental data, reported manufacturing tolerances or expert opinion.

$$\boldsymbol{x} \sim \mathcal{H} \tag{2}$$

Often, for simplicity, a named distribution is assigned to each of these parameters Ilas and Liljenfeldt [16]; Briggs et al. [17]. On top of this, covariances can be considered between parameters which may capture some tendency for groups of input parameters to be more similar or dissimilar. Using the determined statistical behavior of the input parameters, sets of input parameters should be sampled ($\boldsymbol{x}_i$) a number of times (N) and carried through the model to get some code-calculated result y_i, this is mathematically written in Equation (3).

$$f(\boldsymbol{x}_i) = y_i \quad i = 1, 2, \ldots, N \tag{3}$$

If a sufficient number of samples are calculated, it is expected that the sample distribution of y_i will converge. Conclusions can then be drawn about the range of likely results, given the uncertainty distributions defined for the input parameters. Common metrics for analysis on the set of y_i, $i = 1, 2, \ldots, N$ include the mean ($\bar{y}$) and variance (σ_y^2), shown in Equation (4) and (5), respectively.

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{4}$$

$$\sigma_y^2 = \frac{1}{N}\sum_{i=1}^{N} (y_i - \bar{y})^2 \tag{5}$$

The total Monte Carlo method for uncertainty quantification offers the advantage of being applicable to almost any computational model. With a sufficient number of samples, nonlinear or complex relationships between inputs and outputs can be completely captured because no reduced-order models are used. Furthermore, the large number of samples needed to obtain a converged distribution for the output can be used for other analyses (as is done in the current study). The main drawback of this method is its computational cost, depending on the number of uncertain parameters in the input and the nature of the calculation, this method can become infeasible. Also, it is important that this method be performed with separate sensitivity methods such as those described in Section 3.4. Finally, in projects which seek to increase the accuracy of their analysis by iteratively performing UQ/SA on the outputs and then improving input parameter uncertainty estimations with focused efforts informed by the UQ/SA results, a completely new set of calculations must be run for each new set of parameter uncertainty estimations. To explain, if experiments are carried out and a certain set of uncertainties are observed, some amount of computational time can be dedicated to propagating these uncertainties through a model. Then, the most influential sources of uncertainty can be identified and more experiments can be performed to reduce the uncertainty in those improtant input parameters. In the case of total Monte Carlo method, a while new set of models would need to be run in order to propagate the new uncertainty set. This is not the case for the method described in Section 3.2.

3.2. Algebraic UQ using surrogate model

The next method for uncertainty quantification that will be discussed is algebraic uncertainty propagation using a surrogate model. This method essentially consists of two steps: 1) creating and evaluating the validity of a linear model trained on an input domain which encompasses the uncertainty bounds for all input parameters and 2) using random variable algebra to propagate input-specific uncertainty into the output response of interest.

As mentioned, the first step is to create and evaluate a linear model. Any general linear model can be expressed as given in Equation (6). Here, $\widehat{f}$ denotes the linear approximation to the full model f and $\widehat{\boldsymbol{\beta}}$ denotes a vector of fitted linear model parameters in $\mathbb{R}^{N+1}$ where N is the number of uncertain input parameters. Here, a constant 1 is prepended to $\boldsymbol{x}$ such that the translation parameter (often referred to as $\widehat{\beta}_0$) can be contained in $\widehat{\boldsymbol{\beta}}$.

$$\widehat{f}(x) = \widehat{y} = \boldsymbol{x}^T\widehat{\boldsymbol{\beta}} \tag{6}$$

From here, it is necessary to generate a data set of model inputs and full model evaluations that will be used to train or evaluate the linear model. In this study, the same samples used in total Monte Carlo uncertainty propagation are used to create the linear models. Following this, an overdetermined system of equations can be constructed using M pairs of input and code calculated responses as shown in Equation (7). Finally, the normal equations can be used to determine $\widehat{\boldsymbol{\beta}}$ that minimizes least square error between linear model predicted results, $\widehat{f}(\boldsymbol{x})$ and code calculated result $f(x)$.

$$\begin{bmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \\ \vdots \\ \boldsymbol{x}_M^T \end{bmatrix} \widehat{\boldsymbol{\beta}} = \begin{bmatrix} f(\boldsymbol{x}_1) \\ f(\boldsymbol{x}_2) \\ \vdots \\ f(\boldsymbol{x}_M) \end{bmatrix} \tag{7}$$

Equation (8) shows the solution to the normal equations where $\boldsymbol{X}$ is the $M \times N$ matrix containing sampled input parameters and $\boldsymbol{y}$ is the vector of code calculated results. Using this $\widehat{\beta}$, the linear model can be evaluated on a set of sampled inputs and code calculated results.

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} \tag{8}$$

Following the formation of the linear model, algebra of random variables can now be used to propagate input parameter uncertainty. First, as explained in the previous section, some probability distribution must be assigned to $\boldsymbol{x}$ which is chosen to reflect the uncertainty in each of the input parameters. First, the expected value of $\hat{y}$ can be quantified, as is shown in Equation (9). This equation is comparable to Equation (4) in the previous section.

$$\mathrm{E}(\widehat{y}) = \overline{y} = \mathrm{E}(\boldsymbol{x})^T\widehat{\boldsymbol{\beta}} \tag{9}$$

From here, Equation (6) can be rewritten in terms of its covariance as shown in Equation (10), given that $\hat{y}$ is a scalar result. By expanding $\mathrm{Cov}(\boldsymbol{x}^T\widehat{\boldsymbol{\beta}})$ in terms of expected values, the relation given in Equation (11) can be obtained. This result is directly comparable to Equation (5) in the previous section. Both of these equations indicate the uncertainty in a model prediction due to uncertainties in $\boldsymbol{x}$.

$$\mathrm{Var}(\widehat{y}) = \mathrm{Cov}(\boldsymbol{x}^T\widehat{\boldsymbol{\beta}}) \tag{10}$$

$$\mathrm{Var}(\widehat{y}) = \sigma_{\hat{y}}^2 = \widehat{\boldsymbol{\beta}}^T\mathrm{Cov}(\boldsymbol{x})\widehat{\boldsymbol{\beta}} \tag{11}$$

This method offers the advantage of increased computational efficiency for many application scenarios. It also offers direct information on the contribution of each input parameter to output uncertainty through the fitted linear model parameters. On top of these benefits, once the linear model is created, it is easy to propagate new estimates of input distribution uncertainty through the model (given that these new input distributions lie within the domain used to create the linear model). This can be very useful for continuing projects which seek to improve the accuracy of their models with more accurate estimates for input parameter distributions based on the results of previous UQ/SA. However, the major drawback of this method is the linearity assumption implicit in the creation of the surrogate model. This method may not be suitable when input parameter uncertainties are large enough that the output does not vary linearly with the input. Fortunately, the validity of the surrogate model can be tested to evaluate the strength of UQ using this method. Also, it is still possible for the computational cost of model training and testing to be significant given large numbers of uncertain inputs or expensive computational models. It is also worth mentioning that the linearity in the relationship between a certain set of input parameters and an output does not guarantee linearity between that set and a different output parameter. The relationships between input and outputs should always be evaluated for linearity. Later, in Section 4.2 the results and computational cost of this method will be compared to the total Monte Carlo method.

3.3. Multiple linear regression with uncertain data

In the nuclear reactor modeling and simulation field, it is often difficult to perform SA using a Monte Carlo neutronics codes because the uncertainty associated with calculated results can introduce large errors in many sensitivity methods. The application scenario for this study uses a high-complexity model created for a Monte Carlo code, therefore, it is important to investigate the validity of the results from SA given that there exists MC21 reported uncertainty in k_{eff}. As will be shown, naive application of sensitivity methods on data with Monte Carlo uncertainty can lead to inaccurate results.

The intention of the following formulation is to capture the effect of code-reported result uncertainty on the linear model parameters calculated with the method described in Section 3.2. To capture this, it is helpful to first characterize the vector $\boldsymbol{y}$, shown in Equation (8), as a random variable with some multivariate distribution. Here, a normal distribution will be used to describe the uncertainty in each code calculated $f(\boldsymbol{x_i})$, it will be assumed that the code-calculated outputs are independent with identical variances (this will often be the case for results generated from separate calculations of a Monte Carlo neutronics code). Following this, Equation (12) gives a multivariate distribution with mean $\boldsymbol{\mu_y}$ and covariance $\sigma_{mc}^2\boldsymbol{I}$. Each $\boldsymbol{\mu}_{\boldsymbol{y},i}$ is the code-reported estimate for the response of interest $f(\boldsymbol{x}_i)$. σ_{mc}^2 is the variance reported by the Monte Carlo code for the relevant result and it is multiplied by the identity matrix $\boldsymbol{I}$ to yield a diagonal covariance matrix with identical entries.

$$\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu_y}, \sigma_{mc}^2\boldsymbol{I}) \tag{12}$$

With this distribution defined, Equation (8) can be written in terms of covariances in Equation (13).

$$\mathrm{Cov}(\widehat{\boldsymbol{\beta}}) = \mathrm{Cov}((\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}) \tag{13}$$

After expressing the covariance in terms of the linear expected value operator, some algebra can be performed to obtain the final expression shown in Equation (14), this result can also be found in Faber [18].

$$\mathrm{Cov}(\widehat{\boldsymbol{\beta}}) = \sigma_{mc}^2\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1} \tag{14}$$

These methods will be used to provide uncertainty bounds on the linear model parameters calculated in Section 4.1.

3.4. Linear and nonlinear sensitivity analysis

Four separate SA methods will be used in this study, two linear methods and two nonlinear methods. The first method to be presented is the standardized regression coefficient (SRC) method. Put simply, this method fits a linear model between inputs and outputs after each has undergone a standardization transformation. The standardization transformation is defined for the arbitrary vector $\boldsymbol{s}$ in Equation (15). Here, $\bar{\boldsymbol{s}}$ refers to a vector with a mean of $\boldsymbol{s}$ and standard deviation of σ_s.

$$\boldsymbol{s}' = \frac{\boldsymbol{s} - \bar{\boldsymbol{s}}}{\sigma_s} \tag{15}$$

The standardization transformation is performed on the vector $\boldsymbol{y}$ and column-wise on the design matrix, $\boldsymbol{X}$. Following this, linear regression is performed on the standardized forms of $\boldsymbol{X}$ and $\boldsymbol{y}$ using Equation (8). The resulting $\widehat{\beta}$ yield dimensionless sensitivity coefficients that quantify the contribution of the uncertainty in each

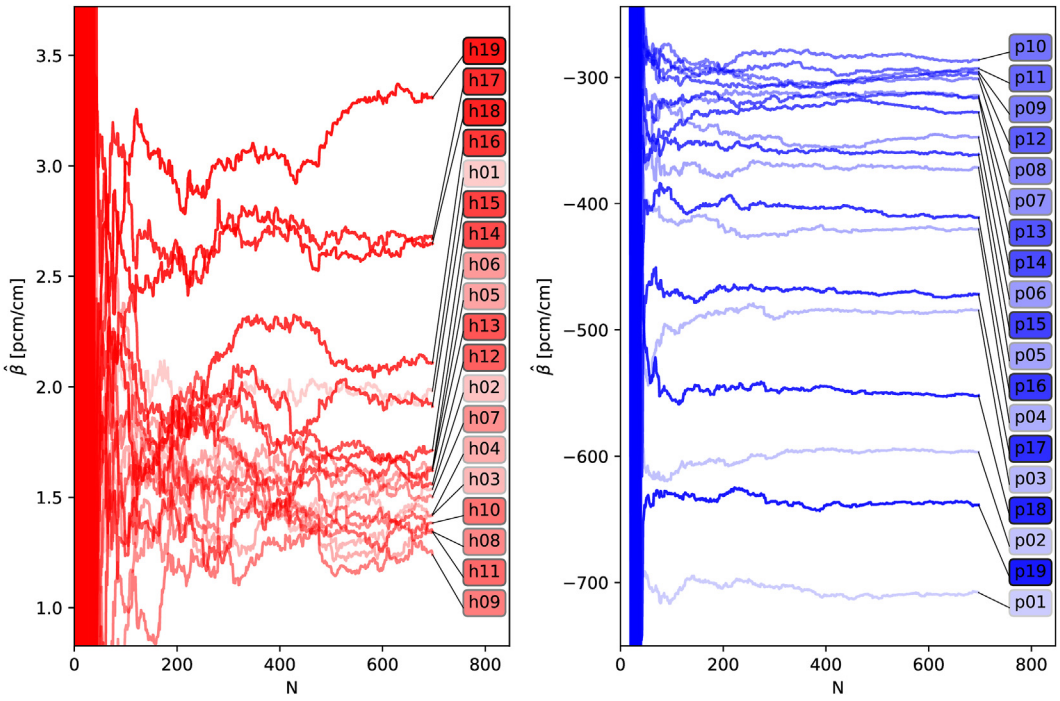

Fig. 4. Estimates for linear model parameters as a function of number of samples used for estimation. The left plot shows the linear model parameters corresponding to the fuel meat height for each of the 19 plates, the right plot shows the linear model parameters corresponding to the inset distance for each of the 19 plates.

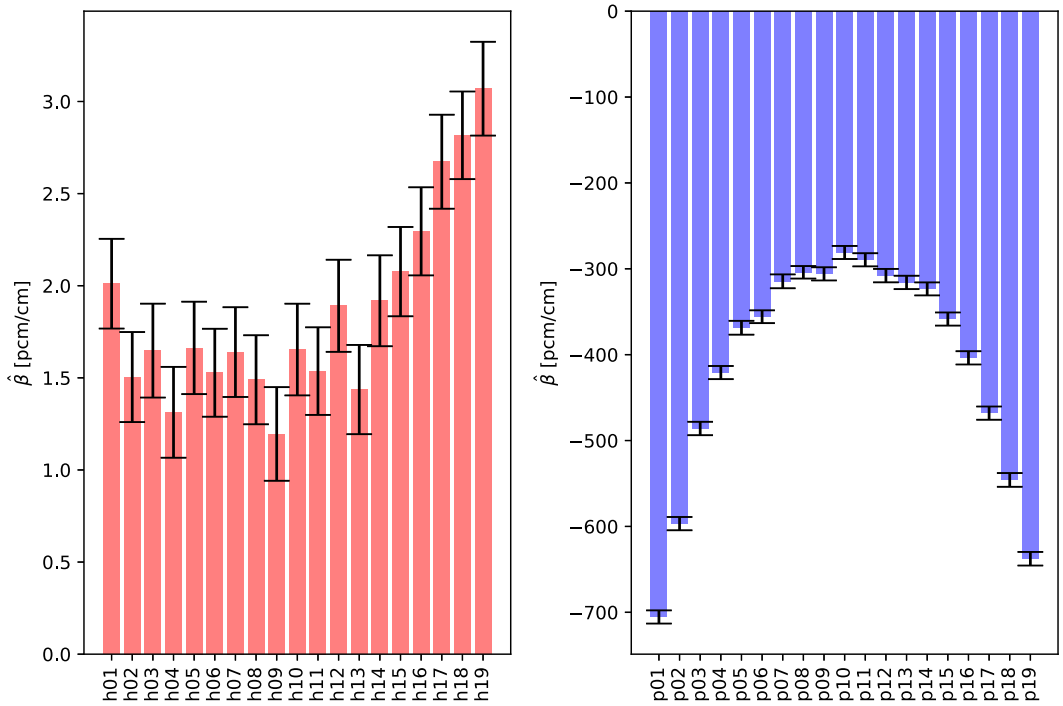

Fig. 5. Values for linear model parameters created using 350 samples. Uncertainties in $\hat{\beta}$ due to code-reported Monte Carlo uncertainty in k_{eff} used to train the model also shown. The left plot shows the linear model parameters corresponding to the fuel meat height for each of the 19 plates, the right plot shows the linear model parameters corresponding to the inset distance for each of the 19 plates. The error bars correspond to the 95% confidence interval for these values.

input parameter to the uncertainty in the scalar output. This method is particularly useful when the input parameters are found to be linearly related to the output. If this is not the case, an alternative version of this, called standardized rank regression coefficient (SRRC) should be used Hamby [19].

SRRC can be used for nonlinear applications because it draws conclusions about the monotonicity of the relationship between an input and output, not necessarily the magnitude of effect (which can be nonconstant over the domain). This method is similar to the SRC method except the rank transformation is used instead of the

standardization transformation. For the rank transformation, each element in $\boldsymbol{s}$ is replaced by an integer which indicates its position if the set of all of $\boldsymbol{s}$ was ordered smallest to largest. Again, this rank transformation is performed on the vector $\boldsymbol{y}$ and column-wise on the design matrix, $\boldsymbol{X}$. Following this, linear regression is performed on the ranked forms of $\boldsymbol{X}$ and $\boldsymbol{y}$ using Equation (8).

The next sensitivity method is the partial correlation coefficient (PCC) method. This method seeks to report a correlation between a particular input variable and the output while controlling for confounding correlations from other input variables, calculation methods for partial correlation coefficients in high-dimensional input space is nontrivial Cramér [20]. This method relies on linearity of the input/output relationship. For this study, a python-based tool for calculation of partial correlation coefficients is used Vallat [21]. For reference, the formula for a 2-D input space is given in Equation (16). Here, y is used to indicate the output of interest and x_1 and x_2 are the two input variables. $r_{a,b}$ denotes the Pearson correlation coefficient between arbitrary variable a and b and $r_{y,x_1|x_2}$ indicates the PCC between x_1 and y when controlling for x_2. The Pearson correlation coefficient is defined in Equation (17).

$$r_{y,x_1|x_2} = \frac{r_{x_1,y} - r_{x_1,x_2} r_{x_2,y}}{\sqrt{(1 - r_{x_1,x_2}^2)(1 - r_{x_2,y}^2)}} \tag{16}$$

$$r_{a,b} = \frac{\sum_{i=1}^{N} (a_i - \overline{a})(b_i - \overline{b})}{\sqrt{\sum_{i=1}^{N} (a_i - \overline{a})^2} \sqrt{\sum_{i=1}^{N} (b_i - \overline{b})^2}} \tag{17}$$

The final method, the partial rank correlation coefficient method (PRCC), involves performing the analysis described in the previous paragraph on the rank-transformed versions of $\boldsymbol{X}$ and $\boldsymbol{y}$.

The sensitivity measures resulting from these methods are useful in characterizing the relationship between the uncertainty in the input parameters to the uncertainty in the final output. When these sensitivity measures are compared across methods, it is important to only compare them on the basis of their relative ranking within the results of a single method and sign. These sensitivity methods are performed for both sets of uncertain parameters in Section 4.3, the input variables with the largest influence on the uncertainty in k_{eff} are identified.

4. Results

4.1. Linear model creation and evaluation

The following subsection will be further divided into the results pertaining to a linear model created for the uncertain fuel geometry parameters and those pertaining to a linear model created for the uncertain fuel composition parameters. Each of these smaller subsections will show the convergence of the linear model parameters as a function of samples run. Following this, some threshold will be determined for convergence and the remaining unused samples will be used for model evaluation. Finally, the methodology described in Section 3.3 will be used to quantify the effect that code-reported Monte Carlo uncertainty has on the linear model parameters of each model.

4.1.1. Fuel geometry parameters

The distributions for the uncertainty in the fuel geometry parameters, as listed in Table 1 are used to generate a set of samples that can be used for model training and evaluation. These distributions were selected for two reasons:

1. These linear models are created to predict reactor characteristics when the parameters fall within the domain of their reported uncertainty (as should be the case for realistic reactor operation). Therefore, by sampling the parameters only in this domain, the model will be optimized to predict reactor behavior in a realistic domain. These linear models are not being used to predict reactor characteristics when the concerned parameters may be deliberately perturbed outside the domain of their uncertainty.
2. On a practical level, this allows for the same calculation set to be used for both total Monte Carlo uncertainty propagation and the creation of the linear models.

The computational workflow depicted in Appendix A. was used to generate a set of sampled input parameter sets and model results.

In total, 698 MC21 inputs were generated using sampled fuel geometry parameters and k_{eff} was calculated for each input. Such a large number of samples was necessary due to the large number of uncertain input parameters considered in this analysis. Fig. 4 shows an estimation for the linear model parameters $\widehat{\beta}$, calculated using Equation (8) as a function of number of samples. This figure is included to show the ranking and convergence of the linear model parameters. Fig. 5 is included to show the uncertainty and value of the linear model parameters with black error bars. In this figure, the linear model parameters corresponding to both the fuel meat heights and inset distances are shown. These results can be viewed as a preliminary sensitivity analysis because the change in the response of interest, k_{eff}, is estimated for some change in a single parameter. As is seen through the sign of the linear model parameters relating to the fuel plate inset distance, increases in the inset distance reduce core criticality, this is expected because by increasing the inset distance, total fuel volume is decreased. The opposite can be said about increasing the fuel meat height, hence the positive sign associated with the linear model parameters relating to these inputs. Based on these results, it is decided that the linear model used for the evaluation of the linearity assumption will be generated using the first 350 samples. Although it appears that the linear model parameters pertaining to the fuel meat height seemingly did not converge even for the full 698 sample set, results shown later in this section will attribute this to the Monte Carlo uncertainty associated with code calculation. Also, the seemingly parabolic shape of the linear model parameters associated with the inset distances is caused by the higher fluxes in the outer fuel plates. With higher fluxes, the dimensions of these plates are more influential on the core k_{eff} than the inner plates. This trend is much weaker in the linear model parameters relating to the fuel meat height due to both the significant uncertainty associated with these quantities as well as the differing plate widths which cause a perturbation in plate 1 result in less added fuel mass than the same perturbation in plate 19. This geometrical effect diminishes the importance of the fuel meat height relating to the lower-numbered fuel plates.

Following this, the uncertainty in the linear model parameters due to Monte Carlo uncertainty in code-calculated k_{eff} will be addressed. Using the formulation given in Section 3.3 an uncertainty can be assigned to each of the linear model parameters that comes about from the Monte Carlo uncertainty in the code-reported k_{eff}. The magnitude of this uncertainty is shown in Fig. 5. The relative uncertainties associated with the fuel meat height parameters are much larger than those associated with the fuel plate inset distance. Earlier, the linear model parameters shown in Fig. 4 did not seem to converge even at the maximum number of samples. This random noise is not an effect of an

Table 3
Convergence of linear model parameters for number of samples run. $\widehat{\beta}_N$ indicates the linear model parameter estimated with N samples for each of the input parameters.

Parameter	$\widehat{\beta}_{10}$	$\widehat{\beta}_{30}$	$\widehat{\beta}_{60}$	$\widehat{\beta}_{100}$	$\widehat{\beta}_{140}$	$\widehat{\beta}_{282}$
Uranium-aluminide mass in element [pcm/g]	0.103	−0.0154[a]	−0.0453	−0.0217	−0.0210	−0.0474
Boron-carbide mass in element [pcm/g]	−1479	−1466[a]	−1463	−1462	−1465	−1462
Aluminum (X8001) mass in element [pcm/g]	0.190	−0.0754[a]	−0.0738	−0.0727	−0.0734	−0.0691
Total uranium mass in element [pcm/g]	−7.08	−7.09*	−6.29	−6.37	−6.34	−6.37
Total uranium-235 mass in element [pcm/g]	37.6	28.3[a]	28.6	28.0	28.1	27.9
Boron-10 abundance [pcm/%]	−406	−397[a]	−397	−397	−397	−398
Uranium-234 fraction [pcm/%]	−268	−209[a]	−163	−164	−179	−213
Uranium-236 fraction [pcm/%]	−319	440[a]	−33.1	−0.576	−29.5	11.6

[a] Indicates calculated $\widehat{\beta}$ is converged within the bounds of its uncertainty. See text for details.

Table 4
Uncertainties in linear model parameters due to Monte Carlo uncertainty in code-calculated k_{eff}.

Parameter	Absolute Uncertainty	Relative Uncertainty
Uranium-aluminide mass in element	0.028 pcm/g	131%
Boron-carbide mass in element	5.18 pcm/g	0.35%
Aluminum (X8001) mass in element	0.016 pcm/g	22%
Total uranium mass in element	0.287 pcm/g	4.53%
Total uranium-235 mass in element	28.1 pcm/g	1.13%
Boron-10 abundance	0.66 pcm/%	0.17%
Uranium-234 fraction	28.4 pcm/%	15.8%
Uranium-236 fraction	64.0 pcm/%	217%

inadequate number of samples and is instead noise that will exist in linear model parameter estimation for any number of samples due to uncertainty in the training set. Put simply, the fuel meat height was not perturbed enough to induce an effect on k_{eff} that is as clearly discernible from the random noise on k_{eff} generated from the Monte Carlo calculation method. Although the sign and general magnitude of these linear model parameters can be seen, a clear ranking of the parameters is not.

Finally, it is important to evaluate the validity of the linear model. Although the linear model will not necessarily be used explicitly to predict k_{eff}, this analysis gives some insight into the true linearity of the relationship between the fuel geometry parameters and k_{eff}. Linear models will be used later in Section 4.2 to propagate parametric uncertainty through the ATRC model. One useful measure of the accuracy is the regression coefficient (R^2), it measures the closeness of fit of the linear model *to the data used to create the model*. For the linear model created using the aforementioned 350 samples, the R^2 value is 0.997. Furthermore, when using this linear model to predict the remaining 348 samples, a mean absolute error of 3.52 pcm was observed. From these two metrics, it seems that the response of k_{eff} to perturbations in the inset distance and fuel meat height are largely linear in the domain explored in this study.

4.1.2. Fuel composition parameters

The distributions for the uncertainty in the fuel composition, as

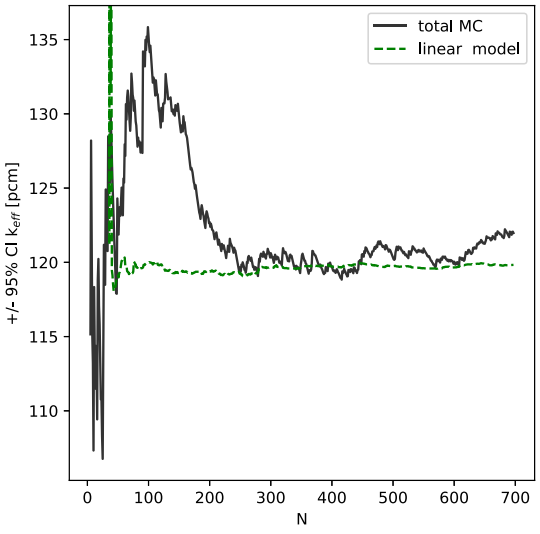

Fig. 6. 95% confidence interval estimate for uncertainty in k_{eff} due to parametric uncertainty in fuel geometry parameters as a function of number of samples used to generate the estimate. Uncertainty propagation methodologies used to generates these estimates are described in Section 3.1 and Section 3.2.

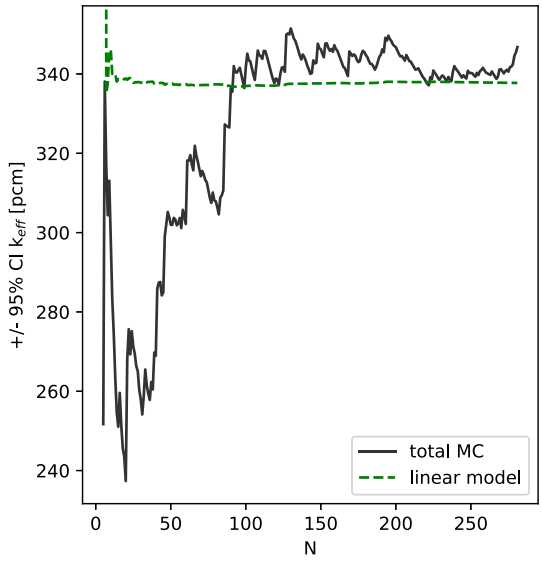

Fig. 7. 95% confidence interval estimate for uncertainty in k_{eff} due to parametric uncertainty in fuel composition parameters as a function of number of samples used to generate the estimate. Uncertainty propagation methodologies used to generates these estimates are described in Section 3.1 and Section 3.2.

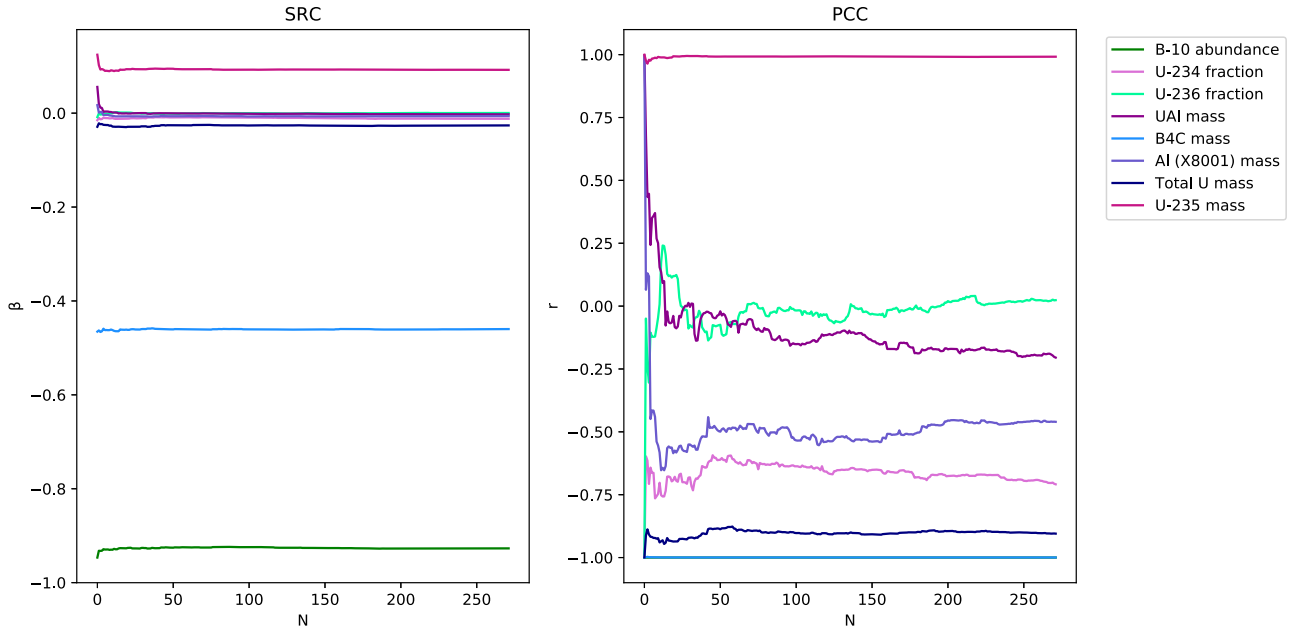

Fig. 8. Estimates for SRC and PCC sensitivity measures for increasing number of samples for uncertain parameters pertaining to fuel composition.

listed in Table 2 are used to generate a set of samples that can be used for linear model training and evaluation. For these parameters, 282 calculations in total were performed to obtain sets of sampled input parameters and values of k_{eff}. This set of parameters is different from the fuel geometry parameters because these parameters are more distinct from each other (in unit and physical representation). In the previous example, the calculated linear model coefficients could be compared—within the same parameter across plates. For the fuel composition parameters, this is not the case because each parameter has a significantly different meaning and often different units. This demonstrates the need for more advanced sensitivity methods to draw conclusions about the relative importance of parameters, as will be shown in Section 4.3.

First, the convergence of the linear model parameters will be discussed. Due to the large differences in magnitude for the assorted linear model parameters, the convergence must be shown in tabular form in Table 3. However, if the reader wishes to see the convergence in a pictorial form, Fig. 8 shows the convergence of the standardized form of these parameters. Nevertheless, all estimates for the linear model parameters converge long before the 140 samples that will be used later to evaluate the linearity assumption. The convergence criteria is not as simple as finding the number of samples needed for the change in estimated $\widehat{\beta}$ to be reasonably small, this is because the Monte Carlo uncertainty in the MC21 results causes there to be constant variability in these estimated parameters. With the fuel geometry set of parameters, there was a smaller uncertainty in the linear model parameters than is the case in the current fuel composition set and also more samples were run so convergence was easier to identify. Therefore, for this set, it is important to establish a convergence criteria that takes into account sample-size independent variability in the estimation. If $\widehat{\beta}_N$ is the estimate for the linear model parameter made with N samples, it will be considered converged if $\widehat{\beta}_{N-1}$ and $\widehat{\beta}_{N+1}$ are within the 95% confidence interval of $\widehat{\beta}_N$, as determined using the uncertainty calculated given in Equation (14). From this, a sample size of 140 can be determined as sufficient to create a linear model with converged parameters. These coefficients can be interpreted as the change in k_{eff} for a perturbation to a particular input parameter while keeping all the other parameters fixed. Hence, the very small reactivity effect in "Uranium-aluminide mass in element" can be attributed to the fact that uranium-235 content in the fuel has been fixed by another parameter in the set.

Table 4 gives the 95% confidence interval uncertainty of the 8 linear model parameters, calculated according to the methodology described in Section 3.3, using 140 samples. From the results, it is clear that even though the uncertainty in k_{eff} was kept below 2 pcm for the MC21 calculations, there can be large uncertainties associated with computed linear model parameters. This is due to the combination of their small parametric uncertainty and small effect on calculated k_{eff}. Therefore, for parameters with unreasonably large uncertainty in their computed coefficients, it is required that perturbations cause parameters to lie outside the bounds of their uncertainty in order to obtain accurate results. As will be seen in Section 4.3, these are also likely to be parameters whose parametric uncertainty has the smallest effect on k_{eff}.

Lastly, it is important to evaluate the validity of the linear model. This model will be used in the next section, Section 4.2, to propagate parametric uncertainty through the ATRC model. Although the uncertainty in the linear model parameters was reported, this is not a complete description of the uncertainty associated with results calculated using the linear model. The R^2 value associated with the linear model is 0.999. To test whether this linearity assumption is valid for perturbations to the input parameters within domains of their respective uncertainty, this linear model is used to estimate the k_{eff} of the remaining 142 samples. The resulting mean absolute error was 1.51 pcm, therefore, it can be determined that the linear model has a high degree of accuracy when predicting k_{eff} of model configurations whose parameters lie within the described domains of uncertainty.

4.2. Comparison of uncertainty propagation methods

Sections 3.1 and 3.2 describe two methodologies for propagating uncertainty through complex models. The first, called the "total Monte Carlo method," repeatedly samples from distributions assigned to each input parameter of interest. These distributions are carefully selected to reflect the uncertainty associated with that

Table 5
Sensitivity measures for five most influential fuel geometry parameters on k_{eff} uncertainty.

Parameter	SRC	PCC	SRRC	PRCC
Inset distance of plate 1	−0.336	−0.985	−0.346	−0.785
Inset distance of plate 19	−0.323	−0.981	−0.305	−0.743
Inset distance of plate 2	−0.304	−0.979	−0.313	−0.754
Inset distance of plate 18	−0.275	−0.974	−0.269	−0.701
Inset distance of plate 17	−0.247	−0.967	−0.252	−0.674

parameter. Then, calculations are performed on many sets of sampled parameters to get some distribution for the response of interest. This distribution in the response of interest treated across the number of samples can be said to reflect the expected uncertainty associated with that calculated result. The second uses a linear surrogate model to algebraically propagate uncertainty. For this method, a domain of interest is identified that encompasses the expected uncertainty for each parameter. It is important that this domain is large enough to encompass the uncertainty of the relevant parameters such that extrapolation is not performed if using the model to analyze the full range of uncertainty. However, the larger the domain becomes, the relationship between input and output becomes less likely to be linear. Then, a training set can be created by sampling from this domain and performing full calculations on the samples. Linear regression can then be performed on the input/output pairs to get a linear model that can be used to algebraically propagate uncertainty according to Equation (11).

The estimated uncertainty in k_{eff} due to the fuel geometry parameters, as a function of number of samples used to generate the estimate, is shown in Fig. 6. The equivalent plot for the fuel composition parameters is shown in Fig. 7. In these figures, x-axis for the uncertainties algebraically calculated with the linear model indicate the number of samples used to train the model. From these results, it is clear that propagating uncertainty using the linear model provides a more accurate estimate of uncertainty for small numbers of samples. This can be extremely useful in reducing the computational cost of uncertainty propagation for models with large evaluation times. Also, if the parametric uncertainties listed in Tables 1 and 2 were reevaluated, these new uncertainties can be trivially propagated through the linear model given that the reevaluated uncertainties are within the domain of sample space used to train the linear model. However, it is important that the validity of the linear relationship between inputs and outputs is evaluated as was done in Section 4.1. The total Monte Carlo method may be more useful for predicting the uncertainty in a model output when there are highly nonlinear relationships between inputs and outputs. Given the high accuracy of the linear model reported in the previous section, total Monte Carlo propagation may not be an efficient method to perform uncertainty quantification for the parameters explored.

Furthermore, due to the higher dimensionality of the input space associated with the fuel geometry parameters, it takes longer for the estimated uncertainty to converge for both methods used as compared to the time it takes for the fuel composition parameters to converge. This further demonstrates the usefulness of using linear regression to create an inexpensive model to propagate uncertainty. In both figures, there is noise in the estimate for the uncertainty even at large values of N. For both parameter sets explored, this noise is on the order of 3 or 4 pcm which is assumed to have a negligible effect on the results. In the end, the uncertainty in the fuel geometry parameters had a smaller total effect on k_{eff} than the uncertainty in the fuel composition parameters. Among other possibilities, one contributing factor for this could be that the fuel plate geometric parameters were perturbed independently by plate. This means that it is very unlikely for each plate to be simultaneously perturbed in the same direction, this leads to lower total deviation from the nominal plate dimensions than if all plates were perturbed simultaneously. In the case of the fuel composition, the plates *were* perturbed simultaneously. This may cause a compounding effect for each plate perturbed and lead to a larger overall effect on k_{eff}.

Table 6
Sensitivity measures for three most influential fuel composition parameters on k_{eff} uncertainty.

Parameter	SRC	PCC	SRRC	PRCC
Boron-10 abundance	−0.927	−0.999	−0.916	−0.981
Boron-carbide mass in element	−0.460	−0.999	−0.433	−0.922
Total uranium-235 mass in element	0.092	0.992	0.085	0.428

4.3. Sensitivity method results

The final set of results presented in this paper relate the uncertainty in the various input parameters to the total uncertainty observed in the output parameter of interest (here, k_{eff}). Four different methods are used, each described in Section 3.4. The SRC method performs linear regression on the input parameters sets and output after each has undergone a standardization transformation. By rescaling the input parameters, the effect on the response of interest can be quantified relative to the range of their uncertainty. To explain, given two parameters of equal influence prior to rescaling (equal β parameters), where one of these two parameters has a larger uncertainty bound, it can be said that the parameter with the larger uncertainty bound contributes more to the output uncertainty than the parameter with the smaller uncertainty bound. After the standardization transformation has been performed, the distributions of both these parameters is identical but the parameter with the larger uncertainty bound has undergone a dilation more significant than its counterpart. Therefore, a unit perturbation to each of these transformed parameters separately will register as a more significant effect for the parameter whose bound was more significantly dilated. The end result is that the SRC method collapses the competing effects of range of uncertainty and magnitude of influence on output into a single metric. The next method is the PCC method, this method finds correlation between an input parameter and output parameter while accounting for confounding correlations with other parameters. As a correlation coefficient, it indicates the explained total variance (not just uncertainty) in k_{eff} that can be attributed to a particular input parameter. Following these two methods are two nonlinear methods, SRRC and PRCC, which operate on the rank transformed versions of the data. Across the four methods, it is useful to compare the signs of the computed sensitivity measures as well as the rankings. *Comparing the numerical values of the sensitivity measures across methods will not yield accurate conclusions.*

Table 5 shows the five most sensitive fuel geometry parameters and their corresponding sensitivity measures for the four methods. All 5 of these parameters pertain to the fuel inset distance, none involve the fuel meat length. Also, the outer plates are ranked as the most influential. This is due to the higher neutron fluxes in these plates. The sign and ranking of these parameters across the four methods indicate strong agreement. The only difference in ranking is the inset distance of plate 19 and inset distance of plate 2 are switched for the nonlinear methods. This consistency across methods indicates that these results can be used to accurately identify input parameters whose individual uncertainty has the

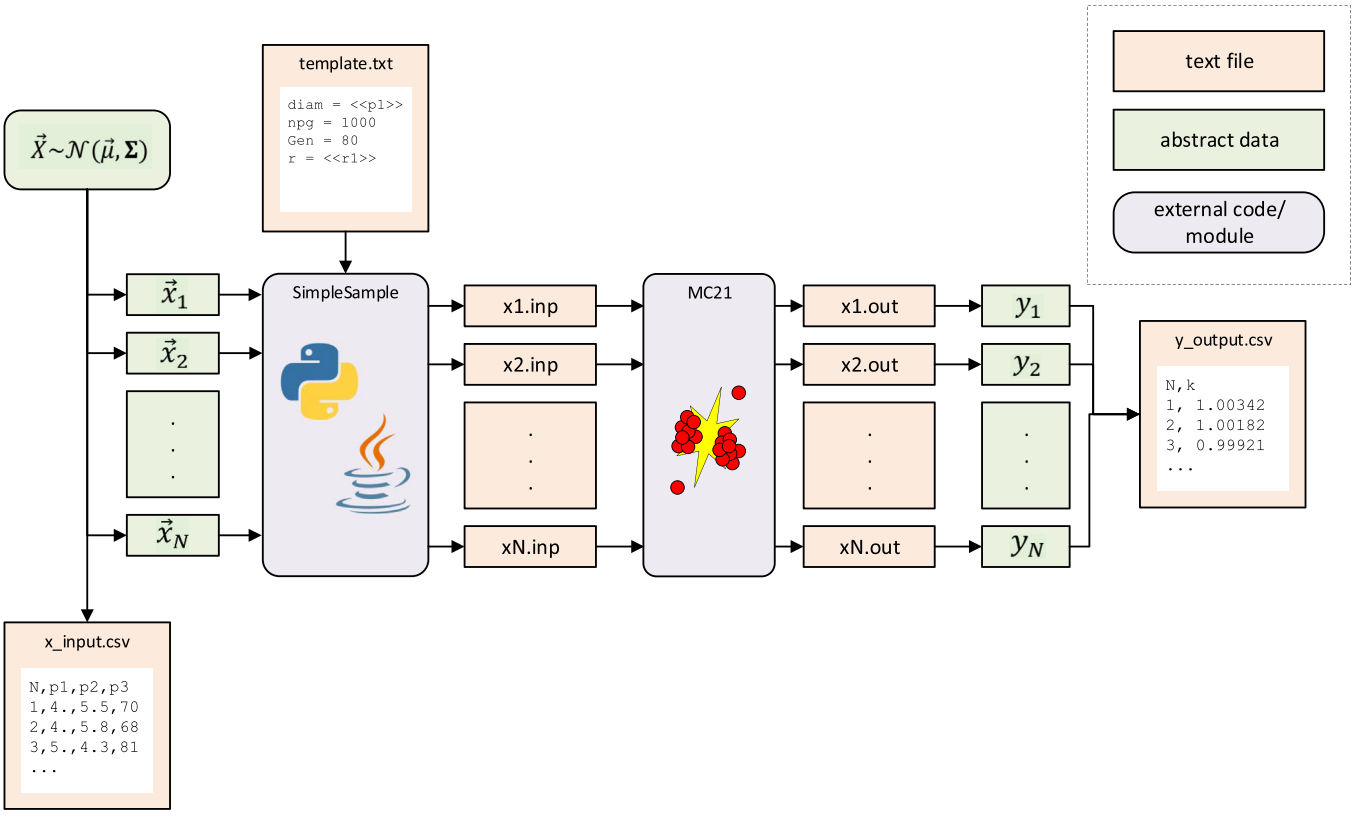

Fig. A.1. Computational workflow used in this study.

largest impact on the overall uncertainty of k_{eff}. In addition, it is helpful to see that the sensitivity measures decrease gradually across these 5 parameters. This demonstrates that efforts should be taken to reduce the uncertainty as many parameters a possible in order to increase the accuracy of k_{eff} predictions because there is no clear cut-off among the 5 parameters shown.

Table 6 shows the sensitivity results for the three most sensitive parameters in the fuel composition parameter set. Fig. 8 shows the convergence of the SRC and PCC measures for this set because a convergence figure for the linear model parameters was not given in Section 4.1.2. Again, the ranking and sign of these sensitivity measures agree across the four methods used. These sensitivity measures are more useful than the linear model parameters included in Table 3 in determining input parameters with the largest effect on k_{eff} uncertainty because of the transformations applied to the data set. In this nondimensionalized form, direct comparisons can be made between parameters within each method. For example, these results indicate that the single largest contributor to output uncertainty by a significant margin is the B-10 abundance. The SRC method indicates the nondimensionalized slope of the linear relationship between input and output along a particular dimension. The boron-10 abundance has twice the SRC measure than the next following parameter, this gives a strong suggestion to scientists and engineers that the uncertainty in this parameter should be reduced in order to obtain more accurate simulation results. This type of analysis is very useful in situations where performing the measurements to reduce parametric uncertainty can be very expensive because it can clearly identify the most important parameters to be investigated. As well, the three most sensitive parameters listed are also the three parameters with the smallest uncertainty in the linear model parameters due to the Monte Carlo error. This is an intrinsic benefit to this methodology, the parameters identified with the largest influence on output uncertainty are also those least affected by random error in the calculated results. Qualitatively explained, this is because these parameters have the largest effect on k_{eff} when considering both the range of their uncertainty bounds and the physical relationship between the parameter and k_{eff}. Therefore the "signal-to-noise" ratio when analyzing these parameters is much better.

5. Conclusions

This paper has provided a detailed methodology for performing SA and UQ on high fidelity models which require Monte Carlo calculation techniques to obtain reactor characteristics. Often, SA in particular is difficult to perform when results are subject to uncertainty from the Monte Carlo calculation techniques because small effects on k_{eff} can be lost in the stochastic noise of the result. This paper has identified a technique to algebraically evaluate the validity of computed linear sensitivities. Furthermore, the results from the dimensionless sensitivity analysis identify the parameters whose uncertainty contribute most to the overall uncertainty associated with k_{eff}. The most influential parameters were also those least affected by the Monte Carlo uncertainty. Future work should explore this result and draw connections between the SRC method and the method to quantify the effect of Monte Carlo uncertainty on linear model parameters.

Also, two different UQ methods were used. Due to the computational resources made available for this study, a method which used algebraic uncertainty propagation through a linear model was compared to the total Monte Carlo method for uncertainty quantification. An unfortunate reality of reduced order modelling is that as the need for a reduced order model increases due to the cost of the full calculation, the capability to evaluate the reduced-order

model decreases for the same reason. In this study, the linear model was created and evaluated with a large number of full-code calculations. By comparing the results derived by the less-expensive propagation using the reduced order linear model with the results derived by the expensive total Monte Carlo method, it can be said that the linear model is sufficient in predicting the uncertainty in k_{eff} for the domain of parameters explored. Future work should explore reactor parameters which are less-likely to have linear relationships with k_{eff}. Exploration of these parameters may require more advanced statistical techniques to create efficient data-driven models.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This manuscript has in part been authored by Battelle Energy Alliance,LLC under Contract No. DE-AC07-05ID14517 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the paper for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

Also, this research made use of the resources of the High Performance Computing Center at Idaho National Laboratory, which is supported by the Office of Nuclear Energy of the U.S. Department of Energy and the Nuclear Science User Facilities under Contract No. DE-AC07-05ID14517.

The authors would also like to acknowledge Daniel Kelly and Scott Spychala as well as the Naval Nuclear Laboratory for their help in code/model development and technical support as well as Ryan Little from the ATR facility. We would also like to thank Majdi Radaideh from the Massachusetts Institute of Technology for the valuable discussions on sensitivity analysis methods we had in the early stages of this work.

Appendix A. Computational Workflow

This section will detail the workflow used to carry out this study, it is included to provide an understanding as to how the sampling process was executed. A flow chart is given in Fig. A.1 that goes through the scripting, MC21 calculation and data aggregation. This workflow was mostly carried out with Python scripting. The parts of this chart labeled "abstract data" were typically stored as various data structures in the Python coding language and referenced when needed. The workflow starts in the top-left corner of the chart, where $\vec{X}$ is some set of uncertain input parameters with mean $\vec{\mu}$ and covariance $\mathbf{\Sigma}$. From here, a set of $\vec{x}_i$ random samples can be drawn from the random variable $\vec{X}$, these random samples are recorded in a.csv file for later analysis. Separately from this, a template input file was prepared for the ATRC model. This template consisted of a nominal ATRC model file with input parameters of interest replaced by easily identifiable text strings. This file, along with the sampled $\vec{x}_i$'s were fed into the SimpleSample python module. This module was created specifically for this study, it replaces the identifiable text strings in the template with the values corresponding to a sample. If multiple samples are fed into it, it will produce multiple output files, each one with the identifiable text strings replaced with the sampled values. In this case, the ATRC model was written using a java API so it must be compiled before a final MC21 input file is obtained. From here, MC21 is run on each of the perturbed input files, labeled as "x#.inp" in the diagram, to obtain output files for each sample, labeled as "x#.out" in the diagram. Finally, each response of interest (denoted as y_i) can be extracted from each output file and stored in a.csv file.

References

[1] R.C. Smith, Uncertainty Quantification: Theory, Implementation, and Applications, vol. 12, Siam, 2013.
[2] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global Sensitivity Analysis: the Primer, John Wiley & Sons, 2008.
[3] M.I. Radaideh, T. Kozlowski, Combining simulations and data with deep learning and uncertainty quantification for advanced energy modeling, Int. J. Energy Res. 43 (14) (2019) 7866–7890.
[4] H. Janssen, Monte-Carlo based uncertainty analysis: sampling efficiency and sampling convergence, Reliab. Eng. Syst. Saf. 109 (2013) 123–132.
[5] D. Rochman, S. van der Marck, A. Koning, H. Sjöstrand, W. Zwermann, Uncertainty propagation with fast Monte Carlo techniques, Nucl. Data Sheets 118 (2014) 367–369.
[6] B.T. Rearden, M.L. Williams, M.A. Jessee, D.E. Mueller, D.A. Wiarda, Sensitivity and uncertainty analysis capabilities and data in SCALE, Nucl. Technol. 174 (2) (2011) 236–288.
[7] D. Price, M.I. Radaideh, D. O'Grady, T. Kozlowski, Advanced BWR criticality safety part II: cask criticality, burnup credit, sensitivity, and uncertainty analyses, Prog. Nucl. Energy 115 (2019) 126–139.
[8] J.C. Helton, J.D. Johnson, C.J. Sallaberry, C.B. Storlie, Survey of sampling-based methods for uncertainty and sensitivity analysis, Reliab. Eng. Syst. Saf. 91 (10–11) (2006) 1175–1209.
[9] M.I. Radaideh, M.I. Radaideh, Efficient analysis of parametric sensitivity and uncertainty of fuel cell models with application to SOFC, Int. J. Energy Res. 44 (4) (2020) 2517–2534.
[10] T. Sutton, T. Donovan, T. Trumbull, P. Dobreff, E. Caro, D. Griesheimer, L. Tyburski, D. Carpenter, H. Joo, The MC21 Monte Carlo Transport Code, Tech. Rep, Knolls Atomic Power Laboratory (KAPL), Niskayuna, NY, 2007.
[11] M.B. Chadwick, M. Herman, P. Obložinský, M.E. Dunn, Y. Danon, A. Kahler, D.L. Smith, B. Pritychenko, G. Arbanas, R. Arcilla, et al., ENDF/B-VII. 1 nuclear data for science and technology: cross sections, covariances, fission product yields and decay data, Nucl. Data Sheets 112 (12) (2011) 2887–2996.
[12] S. Kim, B. Schnitzler, Advanced Test Reactor: Serpentine Arrangement of Highly Enriched Water-Moderated Uranium-Aluminide Fuel Plates Reflected by Beryllium, Tech. Rep. HEU-MET-THERM-022, Idaho National Laboratory, 2005.
[13] M.I. Radaideh, D. Price, D. O'Grady, T. Kozlowski, Advanced BWR criticality safety part I: model development, model benchmarking, and depletion with uncertainty analysis, Prog. Nucl. Energy 113 (2019) 230–246.
[14] J. Meija, T.B. Coplen, M. Berglund, W.A. Brand, P. De Bièvre, M. Gröning, N.E. Holden, J. Irrgeher, R.D. Loss, T. Walczyk, et al., Atomic weights of the elements 2013 (IUPAC technical report), Pure Appl. Chem. 88 (3) (2016) 265–291.
[15] R.J. McConn, C.J. Gesh, R.T. Pagh, R.A. Rucker, R. Williams III, Compendium of Material Composition Data for Radiation Transport Modeling, Tech. Rep, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2011.
[16] G. Ilas, H. Liljenfeldt, Decay heat uncertainty for BWR used fuel due to modeling and nuclear data uncertainties, Nucl. Eng. Des. 319 (2017) 176–184.
[17] J.B. Briggs, L. Scott, A. Nouri, The international criticality safety benchmark evaluation project, Nucl. Sci. Eng. 145 (1) (2003) 1–10.
[18] N.K.M. Faber, Uncertainty estimation for multivariate regression coefficients, Chemometr. Intell. Lab. Syst. 64 (2) (2002) 169–179.
[19] D. Hamby, A comparison of sensitivity analysis techniques, Health Phys. 68 (2) (1995) 195–204.
[20] H. Cramér, Mathematical Methods of Statistics, vol. 43, Princeton university press, 1999.
[21] R. Vallat, Pingouin: statistics in Python, J. Open Source Software 3 (31) (2018) 1026.