Proceedings of GLOBAL 2011
Makuhari, Japan, Dec. 11-16, 2011
Paper No. 391324

Fuel Cycle Covariance of Plutonium and Americium Separations to Repository Capacity using Information Theoretic Measures

Anthony Scopatz[1], Jun Li[2], Erich Schneider[1], and Man-Sung Yim[3]

[1]*The University of Texas at Austin, Mechanical Engineering, 1 University Station C2200, Austin, Texas, USA, 78712,*

E-mail: scopatz@mail.utexas.edu

[2]*University of North Carolina at Chapel Hill, Environmental Sciences and Engineering, Chapel Hill, NC*

[3]*KAIST Department of Nuclear & Quantum Engineering, 291 Daehak-roq, Yuseong-gu, Daejeon 305-701, Republic of Korea*

ABSTRACT: A light water reactor, fast reactor symbiotic fuel cycle scenario was modeled and parameterized based on thirty independent inputs. Simultaneously and stochastically choosing different values for each of these inputs and performing the associated fuel cycle mass-balance calculation, the fuel cycle itself underwent Monte Carlo simulation. A novel information theoretic metric is postulated as a measure of system-wide covariance. This metric is the coefficient of variation of the set of uncertainty coefficients generated from 2D slices of a 3D contingency table. It is then applied to the fuel cycle, taking fast reactor used fuel plutonium and americium separations as independent variables and the capacity of a fully-loaded tuff repository as the response. This set of parameters is known from prior studies to have a strong covariance. When measured with all 435 other input parameters possible, the fast reactor plutonium and americium separations pair was found to be ranked the second most covariant. This verifies that the coefficient of variation metric captures the desired sensitivity of sensitivity effects in the nuclear fuel cycle.

KEYWORDS: *nuclear fuel cycle, systems analysis, sensitivity study*

I. INTRODUCTION

The goal of this study is to validate a novel information theoretic measure as a metric for covariance in nuclear fuel cycle simulations. To test this new metric, a sample fuel cycle is analyzed. For this cycle, a parameter pair with a previously demonstrated high covariance [5] is chosen. Thus, the new metric is verified against a known result. The validation here is a necessary step prior to future studies which may apply this metric over many fuel cycle scenarios.

To begin, a light water reactor (LWR), fast reactor (FR) symbiotic fuel cycle scenario was modeled and parameterized based on thirty independent inputs (*e.g.* FR burnup, separation efficiencies, repository drift diameter, *etc*). Individual fuel cycle realizations were attained by choosing values for each of these inputs stochastically. By executing many (~105,000) realizations, the fuel cycle itself underwent Monte Carlo simulation. From here, sensitivities to pairs of inputs were determined with respect to a global fuel cycle response, namely the repository capacity [MTHM/Repository].

To perform system analyses of the above fuel cycle, standard statistical metrics (such as the mean response over all runs) may be calculated. However, the typical correlation coefficients imply a linear relationship between input and response. Given the complexity of the system, linearity is not a safe assumption. Instead, a novel metric that does not depend on the functional form of the response was used as substitute for the covariance. However, this new measure is a composite of other fundamental information theoretic measures. The underlying measures that were used include the Shannon entropy, mutual information, the uncertainty coefficient, and further aggregations of these metrics. To capture the covariant effects of the problem, these measures were extended from their typical two-dimensional formulations to three dimensional equivalents.

Furthermore, plutonium and americium separation efficiencies (SEs) provide an intuitive example of the covariance metric when related back to the repository capacity response. Conventional sensitivity studies have shown that both plutonium and americium independently have large effects on the capacity of a tuff repository. The plutonium and americium separation efficiencies exhibit a high covariance because both ^{238}Pu and ^{241}Am are major heat load contributors. If the Pu separation efficiency is much higher than that of Am, for example, the repository capacity will be strongly sensitive to changes in the Am SE but insensitive to Pu SE. The reverse situation is obtained when the Pu separation efficiency is low relative to that of Am. Therefore, the conditional covariance between the output, the repository capacity, and the two SEs is high.

The entropy-based measures reviewed in this validation provide a systematic, functional form independent method

for quantifying these covariances. Knowledge of such covariances is important in identifying high-leverage fuel cycle design variable pairs.

II. METHODOLOGY

1. Parameter Generation & Metrics

Many performance assessment studies of the nuclear fuel cycle are linear sensitivity analyses of base-case perturbations. However, statistical performance measures are applicable to a fuel cycle simulation model that is amenable to stochastic execution under variation of several independent parameters. In such a model, the fuel cycle (described below) was parameterized as a function of many different initial input parameters. Values for each input are then chosen stochastically from a predefined, physically-valid range. By generating many inputs and tabulating the corresponding results from an underlying essential physics fuel cycle model, the fuel cycle undergoes a Monte Carlo simulation. An input-output vector constitutes a single fuel cycle realization. From here, relevant statistical metrics on the set of realizations yield information on how the system as a whole performs. Two independent parameters (x and y) and one response (R) variable are examined.

Basic fuel cycle analysis may be performed using 'standard' statistical metrics, such as the variance of the response over all runs. However, the usual correlation coefficients imply a linear relationship between input and response. Instead, information theoretic measures have proved more valuable to ranking the importance of input parameters in a stochastic system [1]. This is because entropy-based measures are independent of the underlying functional form $R(x,y)$.

Rather than using information theoretic metrics in their typical 2D formulations, extending them to their three dimensional equivalents is required for handling fuel cycle covariances. If 3D metrics are constructed such that they relate two input parameters to a response, then joint sensitivity information may be computed from a single set of stochastically-generated data. Conditional covariances (conceptually a sensitivity of sensitivity) may be measured for two inputs to one response.

2. Nuclear Fuel Cycle Simulation & Parameterization

A fuel cycle simulation package, which incorporates physics-based submodels for reactor burnup [2] and repository performance [3], served as the platform for this analysis. Full fuel cycle execution time, including all transients in a closed cycle scenario, takes between 10-30 seconds on a modern computational platform. For additional information on this package please refer to [4].

The sodium cooled fast burner - light water reactor symbiotic cycle studied here is more fully detailed in [4]. The corresponding flowchart is seen in Figure 1. Although the model computes recycle passes leading up to equilibrium [2], response data (*i.e.* the repository capacity) is measured only after the fuel cycle has achieved its mass-balance equilibrium.

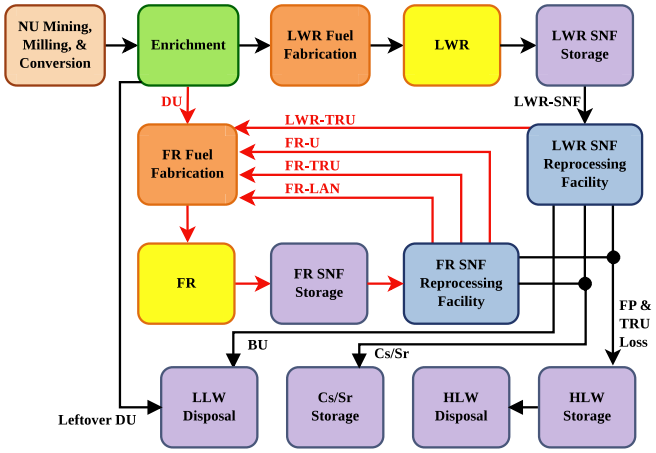

Figure 1: LWR-FR Symbiotic Fuel Cycle Scenario

The fast reactors exist in support of the current and near-future fleet of light-water reactors. The LWR used fuel (UF), after cooling, is sent to an aqueous reprocessing facility. Here the transuranics (TRU) are separated out from the fission products (FP) and the burned uranium (BU). The FP stream is further partitioned at the LWR fuel reprocessing plant. The cesium and strontium are separated out from the remaining FP. The Cs and Sr are then emplaced in a storage facility built for medium-lived isotopes while the BU goes to a low-level waste (LLW) disposal facility. The remaining FPs and any losses from the other streams (TRU, U, Cs/Sr) are treated as high-level waste (HLW) and sent to the repository for permanent disposal.

The recovered transuranics from the LWR are mixed with depleted uranium (DU), and recycled FR uranium and TRU. In this schema, all of the U and TRU discharged from a FR (less reprocessing losses) are recycled into the fast reactor. After being re-burned in the FR, the used fuel once again goes into cooling for three to thirty years. From here the fuel is sent to a FR fuel reprocessing facility. At the FR reprocessing facility the actinide mass of FR-UF is separated from the fission products present. Again, the Cs and Sr are separated from the remaining FP. The actinides are sent back to the fast reactor, the Cs/Sr stream is disposed of in its own storage facility, and the non-Cs/Sr fission products and losses from other streams are sent to the deep geologic repository. After reprocessing but prior to emplacement in the repository, the high level waste stream is cooled for between one and three hundred years. The deep geologic repository itself is modeled after Yucca Mountain.

The fuel cycle just described is then parameterized to be a function of thirty different independent inputs, each chosen randomly on a physically valid range. These stochastic parameters are sampled either in a linearly uniform, log uniform, or one-minus-log uniform fashion. The one-minus-log uniform distribution is useful for choosing separation efficiencies and is known as sampling in the 'nines'. All thirty fuel cycle input parameters and their

sampling functions are defined in more detail in [5].

Thus every fuel cycle realization is its own distinct cycle, tuned for the stochastic parameters selected. However, by bounding the input parameters and comparing many realizations, this study shows that the output parameters are also bounded. The spread of the output parameters due to changes in the inputs determines the relative importance of that input. The change in this spread for one input *x* coming from changes in input *y* represents the covariance information.

3. Statistical Metrics

The statistical structure known as contingency tables was used as a basis for all of the analysis work performed. Such tables have their own detailed history [6-8]. Many metrics have been proposed to measure various quantities about their internal structure. Of these, the class of metrics which is based on *entropy* are perhaps the most useful to sensitivity studies. Many entropy measures are direct parallels to system-wide sensitivities. However, typical formulations involve associating only two parameters to each other. However, the covariances here require a minimum of three parameters: two independent inputs and one response.

Therefore, the 2D expressions are extended to three dimensions. Moreover, to capture covariant effects, rather than joint sensitivities, a novel metric (described below) was implemented. Moreover, 7 bins per axis on the contingency tables were used to obtain acceptable statistical fidelity. For examples of contingency tables with these statistical metrics please refer to [5].

The entropy *H* of a parameter or contingency table is a measure of how evenly spread out the data is over all bins. For contingency tables, maximum entropy implies that all entries in the table have exactly the same value. Conversely, zero entropy implies a fully ordered system. In contingency tables, zero entropy implies that every row and every column have exactly one non-zero entry. All independent input parameters *x* and *y* should have maximum entropy since they are randomly sampled in this experiment.

To compute the entropy, the probability table corresponding to the table is needed. If there are *N* total runs, then any bin in the contingency table may be represented by the matrix element N_{abc}. The subscript *a* indexes the number of *R* response bins *A*, while *b* indexes the number of *x* input parameter bins *B*, and *c* indexes the number of *y* input parameter bins *C*. Thus the probability table may be defined such that:

$$p_{abc} = \frac{N_{abc}}{N}$$

Moreover, marginal sums are represented by a subscript dot notation. The 'dotted' index or indices are the rows or columns that are summed over. The entropy is thusly defined as the sum of $p \cdot \ln(p)$ for any parameter or parameter combination. For instance, the entropy of the response is given as:

$$H(R) = -\sum_{a}^{A} p_{a\cdot\cdot} \ln(p_{a\cdot\cdot})$$

Entropy is not neatly bounded on the range [0,1]. Rather the maximum value for the entropy is given as the natural logarithm of the total number of bins *K* of any table or slice such that the entropy is defined on the range [0, ln(K)].

The next metric considered is the *mutual information I*. The mutual information may be thought of as the overlap in the entropies of various parameters. It is used here to normalize the entropy. As such it indicates how much of a response is determined by a particular input. The mutual information may be calculated in the three-dimensional sense as follows:

$$I(R,x,y) = -\sum_{a,b,c}^{A,B,C} p_{abc} \ln\left(\frac{p_{abc}}{p_{a\cdot\cdot} \cdot p_{\cdot b\cdot} \cdot p_{\cdot\cdot c}}\right)$$

The minimum value of *I* is zero, indicating that *x*, *y*, and *R* share nothing in common. This occurs when all parameters are fully independent. On the other hand, the maximum value of the mutual information depends on the structure of the contingency table and may not be easily expressed in general.

The uncertainty *U* is the metric used as an information theoretic surrogate for a traditional sensitivity. The uncertainty is defined as the mutual information divided by the entropy. The *uncertainty coefficients* are specified with the conditional notation *U(x|R)*. Such coefficients are calculated via [8]:

$$U(x|R) = \frac{I(R,x)}{H(x)}$$

where *H(x)* and *I(R,x)* are the traditional 2D formulations of the entropy and mutual information. This measures the extent to which knowing the input *x* is the same as knowing the response *R*.

In so far as uncertainty coefficients are used in place of sensitivities, an information theoretic replacement for covariance was also found. This type of association is called a sensitivity of sensitivity because it seeks to quantify the sensitivity of *x* given the sensitivity of *y* to the response *R*.

To obtain such a metric, note that each slice of a three-dimensional contingency table is itself a 2D table. It is therefore possible to compute the uncertainty coefficient *U(x|R)* for every slice over *y*. In doing so, a set of *C* uncertainties are generated such that,

$$\{U(x|R)|y\} = \left\{U(x|R)|_{l_0}^{l_1}, U(x|R)|_{l_1}^{l_2}, \ldots, U(x|R)|_{l_{C-1}}^{l_C}\right\}$$

where l is a sequence of $C+1$ points that defines the bin boundaries of y. From here, the mean μ and standard deviation σ of the set may be calculated. However, the choice of slicing along y is arbitrary. One could instead slice along x and compute the set $\{U(y|R)|x\}$ by analogy to the equation above.

Dividing the standard deviation by the mean of these sets yields a *coefficient of variation* c_v:

$$c_v(U(x|R)|y) = \frac{\sigma(U(x|R)|y)}{\mu(U(x|R)|y)}$$

This metric is used as a measure of sensitivity of sensitivities. Unlike the uncertainty coefficient, c_v is not symmetric with respect to x and y. Thus the same pair will have a different covariance rank when measuring with either $c_v(U(x|R)|y)$ or $c_v(U(y|R)|x)$. A symmetric expression of the coefficient of variation is thus obtained by taking the average of the non-symmetric terms:

$$c_v(x|y|R) = \frac{1}{2}[c_v(U(x|R)|y) + c_v(U(y|R)|x)]$$

This measure has the following properties:

1. Defined on the range [0, 1] since $0 \leq U \leq 1$.
2. $c_v(x|y|R) = 0$ implies that both $\sigma = 0$, which indicates that $U(x|R)$ shows no dependence on y. Thus there are no covariant effects observed.
3. $c_v(x|y|R) = 1$ indicates that both $\sigma = \mu$. This connotes that the value of x solely governs the response R from y.

Property 1 is a statement of normalization, which is desirable because it eliminates the need for unfortunate mappings to derive meaning. Properties 2 & 3 are also natural as they provide the useful interpretation of the metric such that high values imply greater covariance. In any sufficiently complex system, such as the nuclear fuel cycle, the extreme values of covariance (0 and 1) are likely impossible. Rather, the relative rankings of $c_v(x|y|R)$ for different (x, y) pairs form the quantitative system analysis tool.

III. RESULTS

The goal of this study, similar to other sensitivity studies [9] or uncertainty analyses [10], is to rank fuel cycle inputs against their effect on the chosen response. The response used here is the total capacity of a fully-loaded repository [MTHM/Repository]. The coefficient of variation $c_v(x|y|R)$ is taken as a measure of covariance effects between x and y. The covariance metric shows that some parameter pairs have a significant joint effect on the fuel cycle. Hence, x and y may exhibit a greater (or less than) commensurate response on the fuel cycle response than the sum of x and y effects individually.

As may be seen in [5], parameter pairs (x, y) are ranked by $c_v(x|y|R)$ for the repository capacity. The top parameters from a 2D sensitivity study all exhibit large effects on the impact of the other thirty inputs examined. Namely, these important inputs include the fast reactor used fuel plutonium separation efficiency (FR_SE_PU), the fast reactor americium separation efficiency (FR_SE_AM), and the high level waste storage time prior to repository emplacement (HLW_Storage_Time).

1. Covariance of Plutonium & Americium Separations

Plutonium and americium separation efficiencies provide an example of the covariance measure. There is already known to be a strong coupling between these two parameters with respect to the repository capacity [11]. Therefore, the coefficient of variation metric should pick this pair as having a high covariance. If it fails to do so, then this provides evidence that the sensitivity of sensitivity measure is not working as intended.

From a 2D sensitivity study [5], both FR_SE_PU and FR_SE_AM are highly ranked inputs individually. They exhibit a high covariance because ^{238}Pu and ^{241}Am are in direct competition for being the top heat load contributor to the repository. Therefore, the degree to which plutonium is recycled in the FR relative to that of americium shifts repository performance in a way that may have a greater impact than when altering FR_SE_PU and FR_SE_AM in tandem.

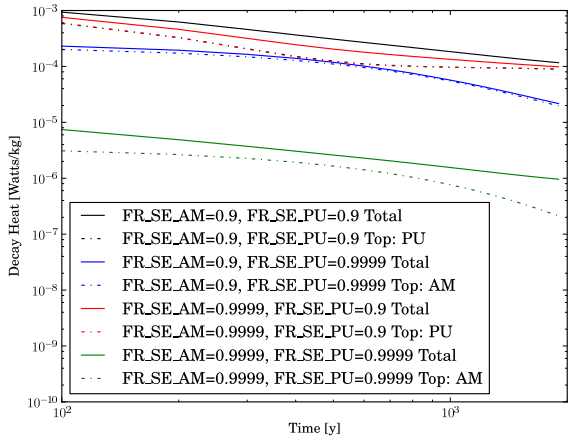

Figure 2: Total & Top Contributor Decay Heat [Watts/kg] of HLW for the parameter pair (FR_SE_AM, FR_SE_PU)

Such covariant effects are visible in Figure 2. Low values of FR_SE_PU and FR_SE_AM have high decay heats and high values have low decay heats. However, for the middle cases where one parameter is high and one is low, the top contributor takes on the value of the low separation efficiency. The switch between Am/Pu and the two order of magnitude range in the decay heat should be captured by the covariance metric.

Computing the coefficient of variation for all $\binom{30}{2} = 435$ parameter pairs possible in this study, the maximal covariance occurs at a value of $c_v(x|y|R) = 0.02318$ for the pair (FR_SE_AM, HLW_Storage_Time). Minimal values are equivalent to zero within error, i.e. $c_v(x|y|R) = 0.0001$. The covariance for the parameter pair of interest (FR_SE_AM, FR_SE_PU) has a value of $c_v(x|y|R) = 0.02316$. This is more than sufficient to give this pair the second rank of all pairs.

Therefore, $c_v(x|y|R)$ satisfies qualitative and quantitative requirements for a sensitivity of sensitivity measure as it successfully determines a high covariance between americium and plutonium separations for the repository capacity response.

IV. CONCLUSIONS

The statistical techniques of contingency table and entropy analysis presented in this paper are widely-used tools for capturing the behavior of complex systems. Here they have been applied to a simulator of the nuclear fuel cycle and used to verify a novel covariance metric. It was shown that the coefficient of variation could illuminate dependencies where changes in one input have a strong effect on the system response to another. This was done by validating the sensitivity of sensitivity metric against a parameter pair which is known to have a strong interdependence.

Therefore, this tool can aid system designers by helping them perceive which design variables jointly matter. Conditioned on a specific system response, like repository capacity, coefficients of variation may guide iterative technology development by demonstrating which design features must be considered together and which can be modified independent of other consideration.

In view of the uncertainties in the direction taken by nuclear energy in the future, coupled with the need to develop technologies now that can support (or even enable) an array of possible outcomes, the metrics presented here could provide invaluable guidance. Information theory provides valuable tools to bound the full range of possible outcomes for any range of inputs. However, the methods must be verified to be physically valid. Such a validation was successfully performed here since the (FR_SE_AM, FR_SE_PU) had was found to be the second most important to the repository capacity of 435 possible parameter pairs. Moreover, its coefficient of variation value was non-trivially far from the error regime. This indicates that this result is statistically significant and not an accidental result of stochastic variation.

REFERENCES

1. H. Liu, W. Chen, A. Sudjianto, “Relative Entropy Based Method for Probabilistic Sensitivity Analysis in Engineering Design,” *Journal of Mechanical Design*, **128**[2], 326 (2006).
2. A. M. Scopatz, E. A. Schneider, “A new method for rapid computation of transient fuel cycle material balances,” *Nuclear Engineering and Design*, **239**[10], 2169-2184 (2009).
3. J. Li, M-S. Yim, D. N. McNelis, “An Efficient and Comprehensive Model for Fuel Cycle System Performance Evaluation,” Proc. GLOBAL 2009, Paris, France, (2009).
4. J. Li, A. M. Scopatz, M-S. Yim, E. A. Schneider, “The Sensitivity of Fuel Cycle Performance to Separation Efficiency,” *Nuclear Engineering and Design*, **240**, 511-523 (2009).
5. J. Li, A. M. Scopatz, M-S. Yim, E. A. Schneider, “Entropy-Based Nuclear Fuel Cycle Sensitivities and Covariances,” *Nuclear Technology*, Submitted for Publication, Sept (2011).
6. B. S. Everitt, “Analysis of Contingency Tables,” Chapman & Hall, 1992.
7. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, “Numerical Recipes 3rd Edition: The Art of Scientific Computing,” *Cambridge University Press*, 2007.
8. Y. Y. Yao, S. Regina, “6 Information-Theoretic Measures for Knowledge Discovery and Data Mining,” in *Entropy measures, maximum entropy principle and emerging applications*, page 115, 2003.
9. A. Scopatz, J. Li, M.-S. Yim, E. A. Schneider, “Nuclear Fuel Cycle Sensitivity to the Variation of Physical Parameters,”. Proc. *PHYSOR 2010*, Pittsburgh, PA, USA, 2010.
10. J. G. Barratt and E. A. Schneider. “Monte Carlo Uncertainty Analysis of a Transmuting Fuel Cycle,” *Transactions - American Nuclear Society*, **90**,249-250, (2004).
11. A. Scopatz, J. Li, M.-S. Yim, E. A. Schneider, “A Systemic Evaluation of the Sensitivity of Fuel Cycle,” Proc. *GLOBAL 2009*, Paris, France, (2009).