INL/CON-24-76308-Revision-0

Creation Synthetic Data to Train a Digital Twin to Predict Reactor Operations

January 2024

Changing the World's Energy Future

Jaden Sonny Palmer, Christopher Ritter, Ryan Hunter Stewart, Ashley Jeanette Brockman Shields, Ben-oni Vainqueur

INL/CON-24-76308-Revision-0

Creation Synthetic Data to Train a Digital Twin to Predict Reactor Operations

Jaden Sonny Palmer, Christopher Ritter, Ryan Hunter Stewart, Ashley Jeanette Brockman Shields, Ben-oni Vainqueur

January 2024

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

Creating Physics-Informed Synthetic Data to Train a Digital Twin for Predicting Reactor Operations

Jaden Palmer†*, Ryan Stewart†, Ashley Shields†, Ben-oni Vainqueur†, Christopher Ritter†

†*Idaho National Laboratory, Street Address/Box Number, City, State Postal Code, jaden.palmer@inl.gov; Ryan.Stewart@inl.gov; ashley.shields@inl.gov; Christopher.Ritter@inl.gov*
**Idaho State Univeristy, 921 S 8th , Pocatello, ID 83209, palmjad2@isu.edu*

[leave space for DOI, which will be inserted by ANS]

INTRODUCTION

Understanding techniques to strengthen the nuclear safeguards regime is crucial in preventing nuclear proliferation due to recent advancements in the nuclear energy industry such as Generation IV reactors and microreactors. Prior to the construction of a nuclear power plant, it is necessary to understand the proliferation potential of the plant's reactor. Digital twins serve as a unique solution to recognizing reactor behavior indicative of nuclear proliferation. A digital twin is defined as a virtual model that works in unison to represent a physical asset, with a transference of data between the virtual and physical assets [1]. This work serves as validation for training a digital twin on synthetic data fabricated via means of Serpent reactor physics and point kinetics equations simulations. In this case this work is based on parameters of Idaho State University's AGN-201 reactor. The synthetic data can then be utilized to train machine learning models in the future to further investigate the utility of these methods. The accuracy of the predicted data is measured against real operational data to verify the reliability of the synthetic data creation methods and decide whether these methods should be used in the future to inform inspectors of a reactor's proliferation potential.

The state of the art in digital twins for autonomous operations of a reactor are limited to describing the mechanical components of the system, as shown in [2]. Multi-physics simulations validated by physical data from mechanical testbeds are employed, but the realities of mechanical physics to reactor physics far contrast. Additionally, to address the problems in detecting proliferation, a safe-guards digital twin (SG-DT) based on a virtual 300MW sodium-cooled fast reactor has previously been developed to both simulate data and train models for the digital twin to recognize these scenarios [3]. The digital twin utilized a virtual testbed based on synthetic data. Without any physical data to verify the projected information, it is uncertain if the digital twin faithfully represents reality.

It is essential to have validated physics models capable of describing the system beyond familiar operation scenarios. Therefore, we propose a synthetic data preprocessing and creation method through reactor physics simulations and data analytics based on parameters of Idaho State University's (ISU's) AGN-201 reactor. Development of this method contributes to the validation of prior work and provides a tool for digital engineering of reactor systems to be employed when machine learning models are not fully capable of describing the system.

AGN-201m

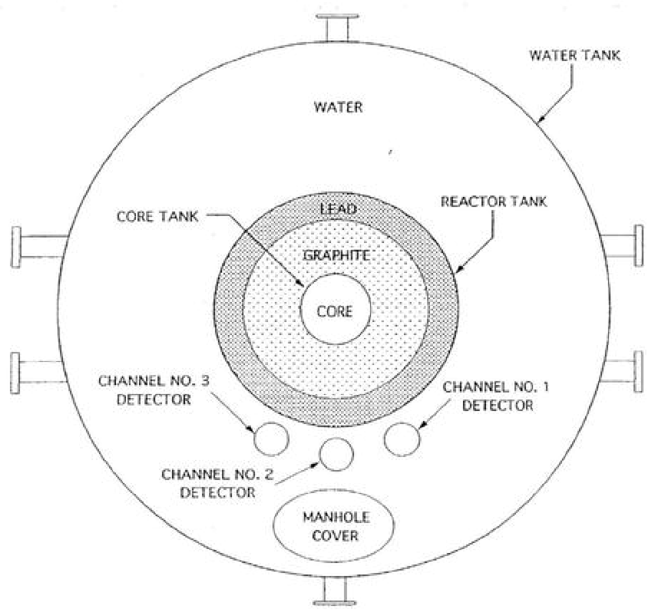

Fig. 1. AGN-201m visual schematic (top view) [4].

The nuclear testbed utilized for this data creation process is Idaho State University's AGN-201m reactor shown in figure 1. The AGN-201m is a small 5W reactor that is employed for teaching and research purposes [4]. Unconventionally, the control rods of the reactor contain the fuel inside of it. The reactor core itself is approximately the size of a football and the control rods consist of 2 safety control rods and a coarse control rod (CCR) that have a worth of $1.69 and a fine control rod (FCR) with a worth of $0.42. Apart from the FCR all other control rods are generally at full insertion into the reactor core (25 cm). During the event of a SCRAM the fueled control rods are instantaneously ejected from the core by gravity and springs.

To measure the reactor's power two channel sensors exist outside of the core with BF_3 filled ionization chambers to measure neutron flux and, subsequently, power. These sensors are known as Channel 2 and Channel 3 and measure power on a logarithmic and linear scale respectively. Both channels also contain the potential to measure reactivity and their respective inverse period. The temperature of the water surrounding the reactor core is monitored as well.

METHODOLOGY

Data Preprocessing

Raw data measured by ANG-201 sensors is stored in DeepLynx, Idaho National Laboratory's data warehouse. To prevent non-physical data being utilized, such as power measured and control rod heights that exceed the AGN-201m's limitations of 5W and 25 cm respectively, a clean function is implemented to replace non-physical data with the respective maximum or minimum value. Intervals in an operation with a high noise to signal ratio may lead to large, short-lived, surge in data. The clean function can recognize these surges and replace them with the average values of their immediate surrounding data.

Critical Data Detection

Nuclear proliferation will not have the capacity to be detected by reactor physics models during the start up or shut down periods. Instead, it will be detected when the reactor is steadily operating at a desired power. This will be when the reactor is critical, generally at 1 mW or above for the ANG-201, with a constant k-effective of approximately 1. For a streamlined approach, we developed an automated critical data detector with capabilities of differentiating critical and non-critical data from operational data of the AGN-201m. To differentiate these two states the critical detector algorithm can recognize:

1. When the fueled control rods enter the reactor core, and the reactor reaches a steady power (critical)
2. When the reactor undergoes a SCRAM and the fueled control rods drop out of the core (shutdown).

To predict startup, the power measured from the Channel 2 detector is utilized. We have broken down the determination of reactor startup is into two distinct phases. The program will first detect when the reactor is supercritical allowing the power increase and will successively recognize when the reactor reaches critical. The criticality detector will sequentially partition all data the occurs in the timestep prior when each event is identified. We use the heights of the FCR and CCR to determine shutdown, since they will immediately drop out of the core after shutdown is initiated. Data after shutdown will be partitioned when the algorithm identifies the heights of the CCR and FCR to be less than 1 cm.

Figure 2 displays predicted K effective (k-eff) values from a Serpent physics surrogate model. The model is trained on reactor characteristics, and creates predictions for

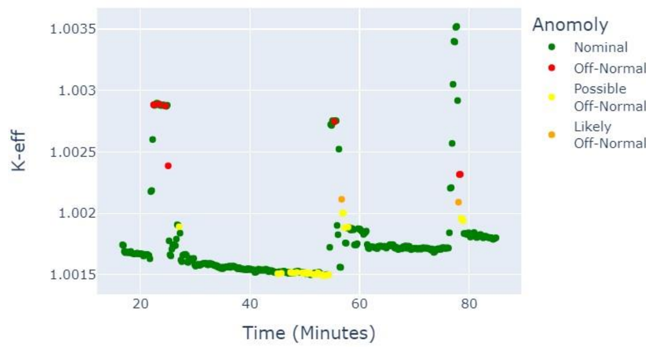

Fig. 2. Predicted K-eff values after criticality detection.

operation based on the water temperature, and control rod heights. Raw data prior to passing through the detection algorithm would lead to initial subcritical values being marked as nominal, and critical data falsely being marked as anomalies and possible proliferation activity.

Point Kinetics Equations Surrogate Model

Coefficients predicted by the previously described Serpent model include k-eff, the delayed neutron fractions (β), decay rates (λ), prompt neutron lifetimes (Λ), and precursor concentrations (C). These coefficients serve as parameters to train a point kinetics equations model (PKE-SM) capable of producing synthetic reactivity and power data. The PKE-SM solves (1) to calculate the reactivity and (2, 3) to calculate the change in neutron density and precursor neutron concentrations. The results from these calculations are then directly transferrable understanding the power produced by the reactor as the change in neutron density will be proportional to the power produced by the reactor.

$$\rho(t) = \frac{k(t)-1}{k(t)} \tag{1}$$

$$\frac{dn}{dt} = \left[\frac{\rho(t)-\beta}{\Lambda}\right] n(t) + \sum_{i=1}^{6} \lambda_i C_i(t) \tag{2}$$

$$\frac{dC_i}{dt} = \frac{\beta_i}{\Lambda} n(t) - \lambda_i C_i(t), \quad i = 1, \ldots, 6 \tag{3}$$

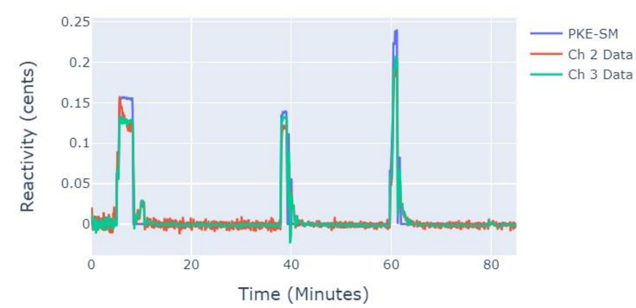

Fig. 3. PKE-SM Reactivity Predictions.

During transient periods, the projected reactivity from the PKE-SM tends to overestimate reality shown in Figure 2.

The prevalence of this overestimation carries a massive impact in predictions for power made by the PKE-SM as (2) takes the area under the curve of the reactivity function to solve the ordinary differential equation and obtain the change in neutron density. These overestimations will not occur outside of transient periods as the PKE-SM takes k-eff to be exactly 1 during critical intervals in the operation. To resolve this issue, we have iteratively found a reactivity correction factor of 0.75 to scale (1), allowing for an accurate approximation compared to physical reality.

Artificial Noise

As power measurements attained via means of neutrons detectors are inherently prone to some level of randomness, I incorporated artificial noise into the synthetic power data. To quantify noise experienced from the physical data I calculated the coefficient of variation (4) over a 500 second period. I then implemented artificial noise into the nonfluctuating synthetic data by randomly selecting values within the Gaussian distribution of the initial data. The Gaussian distribution is calculated by the probability function shown in (5).

$$CV = \frac{\sigma}{\mu} \tag{4}$$

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{5}$$

As the power levels decrease to a threshold below ~15 mW the uncertainty increases due to the low signal that is measured. I address the phenomena by calculating and implementing different coefficients of variation for these different power ranges to create two separate probability densities.

RESULTS

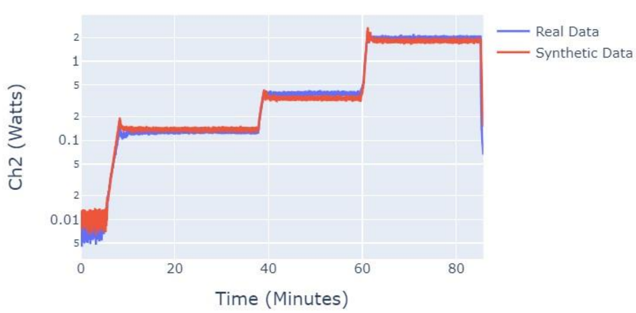

Fig. 2. Final data comparison between synthetic data and actual data during reactor operation.

Figure 3 best qualitatively displays the results of the synthetic data fabrication process. Physical data chosen for comparison is based on power measurements taken from the channel 2 sensor. Comparable results are produced by measurements provided by the Channel 3 sensor as well. Prior to adding the reactivity correction factor the synthetic data captures the general trend relative to physical data. After we integrated the reactivity correction factor and artificial noise it is shown that the synthetic data accurately matches the trend and magnitude of the original data.

To quantitively verify this analysis, the final R^2 score and mean absolute error between both datasets is 0.985 and 0.054 respectively. This high precision further implies the effectiveness of this data creation approach.

CONCLUSIONS AND FUTURE WORK

Although the AGN-201m is small compared to conventional reactors, smaller reactors still exhibit similar behaviors to a large commercial power reactor. The methods outlined provide the general preprocessing and simulation steps to create synthetic data representative of reactor operations. If general operating parameters and physical characteristics of the reactor are provided, then the following steps are transferrable for synthetic data/digital twin creation.

We will build upon these current methods to make the simulation process more convenient and ready for integration. To execute this, we will create another algorithm capable of automatically calculating the reactivity correction factor rather than having to iteratively identify it.

Additionally, a Long Short-Term Memory (LSTM) recurrent neural network will be trained on this synthetic data to further validate the use of these methods for digital twin implementations. Another issue that may be encountered sensor drift, or when the physical data begins to deviate from the derived data over large periods of time due to many factors (i.e. environmental conditions, aging, manufacturing faults, etc).The LSTM neural network will have capabilities to catch this drift, and can signal that reprocessing of synthetic data is necessary from detecting when the error between actual and predicted data reaches a certain threshold.

REFERENCES

[1] M. Tomczyk and H. van der Valk, “Digital Twin Paradigm Shift: The journey of the digital twin definition,” *Proceedings of the 24th International Conference on Enterprise Information Systems*, Apr. 2022. doi:10.5220/0010997600003179

[2] R. Stewart, A. Shields, F. Gleicher, S. Bays, M. Schanfein, G. Reyes, and C. Ritter, “Utilizing a Digital Twin to Analyze and Monitor Proliferation Scenarios in a Sodium-Cooled Fast Reactor,” Journal of Science and Global Security (submitted), 2023.

[3] J. Browning, A. Slaughter, R. Kunz, J. Hansel, B. Rolston, K. Wilsdon, A. Pluth, and D. McCardell, “Foundations for a fission battery digital twin,” Nuclear Technology, vol. 208, no. 7, pp. 1089–1101, 2022.

[4] C. Pope and W. Pheonix, Idaho State University AGN-201 Low Power Teaching Reactor - An Overlooked Gem, ser. Nuclear Reactors Spacecraft Propulsion, Research Reactors, and Reactor Analysis Topics, 2021