

Integral Full Core Multi-Physics PWR Benchmark with Measured Data NEUP 14-6742: Final Report

Shikhar Kumar, Jingang Liang, Miriam Rathbun, Benoit Forget, Kord Smith

April 11, 2018

1 Introduction

In recent years, the importance of modeling and simulation has been highlighted extensively in the DOE research portfolio with concrete examples in nuclear engineering with the CASL and NEAMS programs. These research efforts and similar efforts worldwide aim at the development of high-fidelity multi-physics analysis tools for the simulation of current and next-generation nuclear power reactors. Like all analysis tools, verification and validation is essential to guarantee proper functioning of the software and methods employed. The current approach relies mainly on the validation of single physic phenomena (e.g. critical experiment, flow loops, etc.) and there is a lack of relevant multiphysics benchmark measurements that are necessary to validate high-fidelity methods being developed today.

This work introduces a new multi-cycle full-core Pressurized Water Reactor (PWR) depletion benchmark based on two operational cycles of a commercial nuclear power plant that provides a detailed description of fuel assemblies, burnable absorbers, in-core fission detectors, core loading and re-loading patterns. This benchmark enables analysts to develop extremely detailed reactor core models that can be used for testing and validation of coupled neutron transport, thermal-hydraulics, and fuel isotopic depletion. The benchmark also provides measured reactor data for Hot Zero Power (HZIP) physics tests, boron letdown curves, and three-dimensional in-core flux maps from 58 instrumented assemblies. The benchmark description is now available online and has been used by many groups. However, much work remains to be done on the quantification of uncertainties and modeling sensitivities. This work aims to address these deficiencies and make this benchmark a true non-proprietary international benchmark for the validation of high-fidelity tools.

This report details the BEAVRS uncertainty quantification for the first two cycle of operations and serves as the final report of the project.

2 Benchmark

The BEAVRS benchmark (Benchmark for Evaluation And Validation of Reactor Simulations) represents a 4-loop Westinghouse plant with as much detail available in the open literature. Core loadings and detector readings were provided by plant operators for the first two cycles of operation with all necessary details (i.e. as-fabricated fuel assembly loadings, burnable absorber pin

layouts, operational histories and control rod positions, boron concentration). The core contains 193 Westinghouse optimized fuel assemblies (OFA) each with 264 fuel rods, 24 guide tubes, and one instrument tube arranged in a 17x17 square lattice. First cycle fuel rods in each assembly consist of UO₂ with U-235 enrichments of 1.6, 2.4, or 3.1 weight-percent enrichment, depending on location, as depicted in Figure 1, which also indicates the number of dry annular burnable absorber rods occupying guide tubes. Figure 2 presents the re-load shuffling pattern and fresh fuel loading for cycle 2.

In this reactor plant, there are 58-instrumented assemblies in which one of six variable- mass, in-core fission chambers can be positioned. Each detector passes through a common instrument tube, which is an essential step for the inter-normalization of detector signals. As measurements are taken, detectors are inserted into the core through instrumentation tubes in assemblies until the top of the fuel assembly is reached. Detector signals are then recorded as the detectors are pulled through the core at a constant speed. Thus, measurements represent an integral of signals over a recording interval and this data has been integrated into 61 axial data points for each instrumented assembly location.

Currently, the benchmark document provides detailed geometry for the 3D full core model for all non-proprietary data, and a few geometrical approximations have been made where proprietary data is unavailable. All approximations are documented and are updated as new information is made available. A Python processing script was developed to perform the necessary data filtering and normalization from the fission detectors data. A comprehensive methodology was proposed to quantifying the uncertainty of entire flux map, through estimating the uncertainty introduced in the filtering process and accuracy of detector signals by statistically analyzing the multiple measurements. Both the detailed axial and radial flux maps along with the uncertainties have been obtained and reported previously. For completeness, the process of generating these flux maps will be explained.

The first step in this process of generating radial reaction rate maps is to remove detector background signal. Next, the signal is multiplied by a gain factor to adjust the gain on the detectors. These gain factors are reported for each detector. Figure 3 shows the initial raw detector measurements at Hot Zero Power (HZP), while Figure 4 shows the measurements after correcting for background and applying gain factors.

Figure 4 illustrates that zero points exist where detectors were unable to make measurements. Linear interpolation/extrapolation is used between/from the nearest two points to remove these zero points. The resulting data is shown in Figure 5. Within each detector pass, there are minor fluctuations in core power during the measurement, so the detector signal is divided by power during the measurement. Assembly J10 is the assembly where all detectors make a pass, so this is used for signal normalization purposes. Figure 6 shows the detector measurements in assembly J10 at HZP. It is evident from this plot that each detector has a different signal intensity due to varying Uranium-235 mass in the fission chambers, despite each detector entering the same assembly.

To account for this, each detector pass in each assembly is normalized to the signal strength of that specific detector in assembly J10. This allows all detector signals to have comparable shape functions. Figure 7 shows the detector signals at this stage. There is still some axial misalignment due to varying starting positions during recording, so the signals are subsequently aligned to grid depressions. The positions of these grid spacers are fixed, and Figure 8 presents the detector signals after realignment. The final step is to map the axial measurements from a grid system to a physical axial location on the active fuel. A second order spline fit is used to ensure that there are equal numbers of data points from the Top of Active Fuel (TAF) to the Bottom of Active Fuel (BAF).

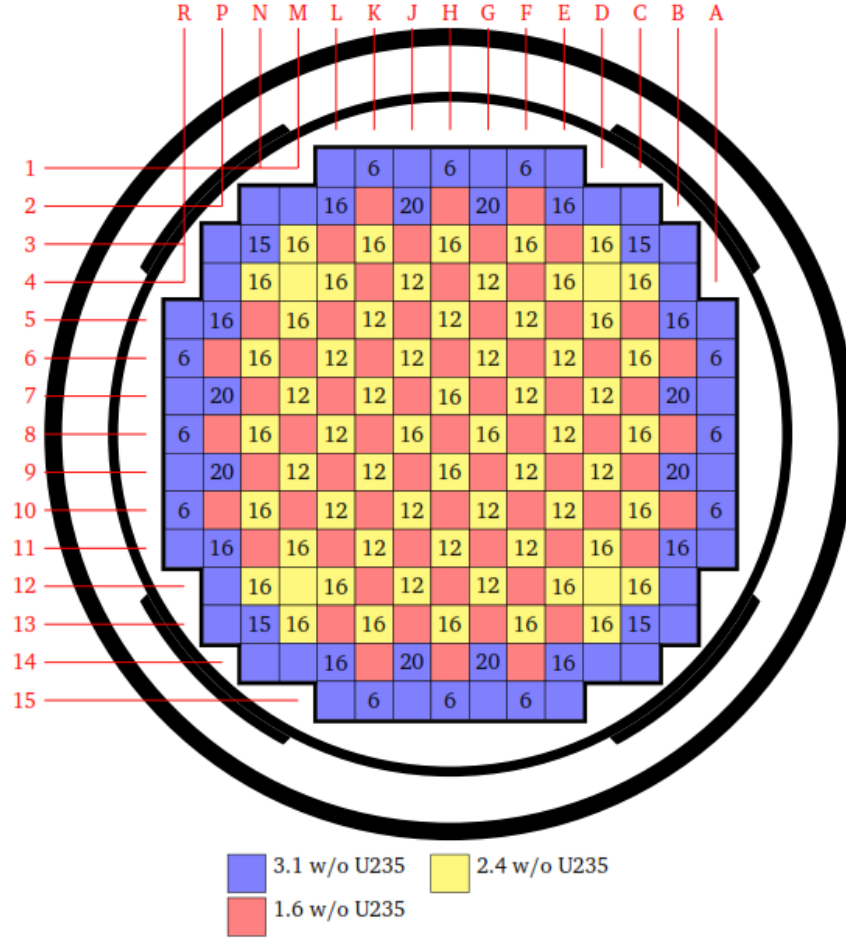


Figure 1: Cycle 1 Loading Pattern

Figure 9 presents this data after the spline fit. The final measurements (no longer normalized to have the same shape functions) are shown in Figure 10.

With this data, axially integrated radial maps of fission reaction rates can now be created. The total sum of all signals in Figure 10 is computed and renormalized to be equal to the total number of assemblies (for this particular case there are 58 assemblies where measurements occurred). Axially averaging each measurement in an assembly yields a relative radial peaking factor, and Figure 11 depicts this factor for each assembly where measurements were taken at HZP. While the specifics of this procedure focus only on data at HZP, the same processing algorithm has been done for all burnups where data is available [1]. It has also been shown that reaction rate maps generated by core simulation software are in close agreement with those shown in Figure 11 [2].

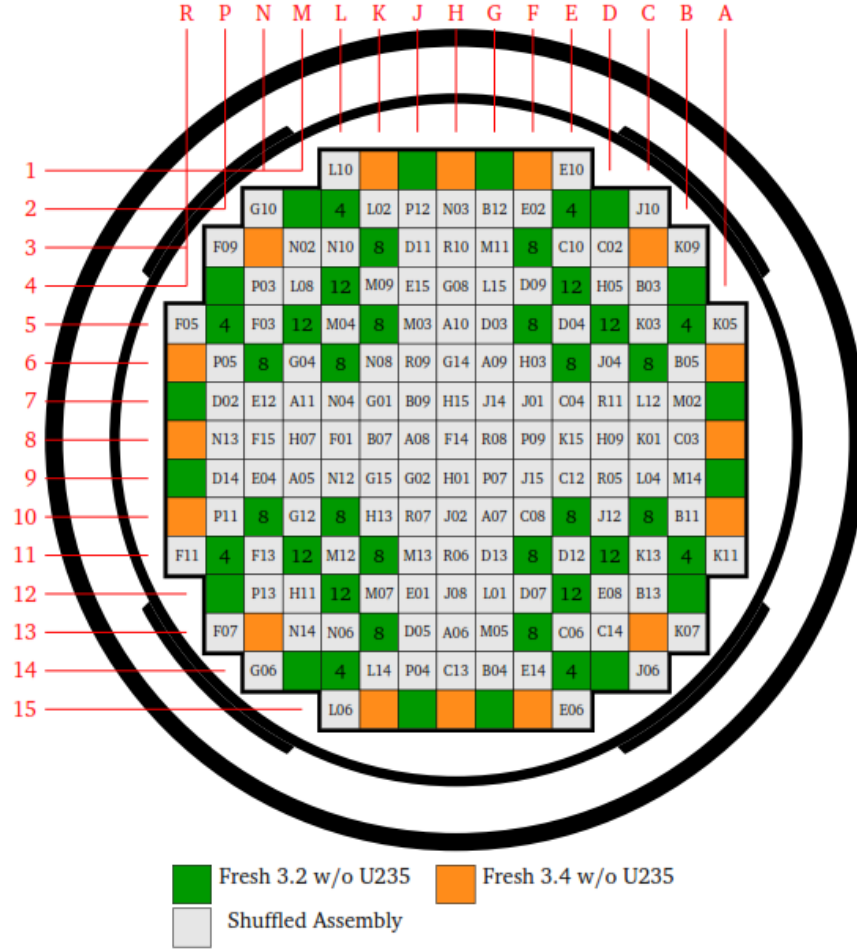


Figure 2: Cycle 2 Loading Pattern

3 Task: Benchmark Documentation

3.1 BEAVRS specifications - Version 2.0.2

The specifications of the benchmark and all data packages are available online at the MIT-CRPG website (crpg.mit.edu). Version 2.0.2 of the BEAVRS benchmark was recently released with a change in the cycle 2 power history plot.

The previous plot was missing the correct End Of Cycle point, at 296 calendar days since BOC, or 10.43 MWd/kg burnup. Figure 13 shows the correct plot and is included in version 2.0.2. Having a complete power history for both cycles is especially important in determining uncertainty through time series analysis, since only full power points will be considered in the analysis, and the Cycle 2 EOC burnup that has been omitted thus far is contained within the subset of full power points.

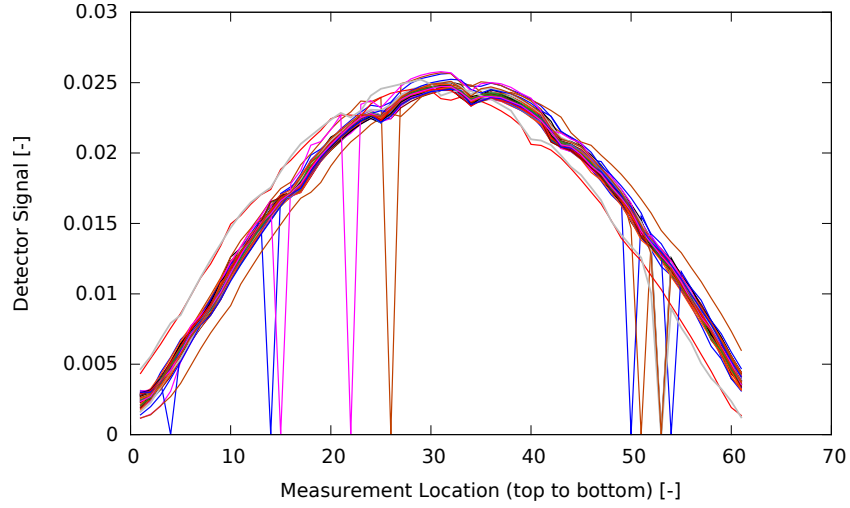


Figure 3: Initial raw detector measurements (top to bottom) [1].

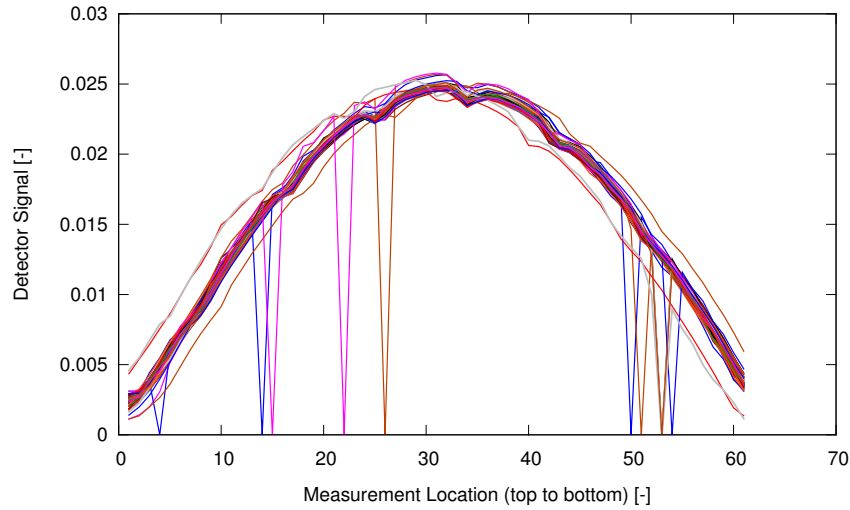


Figure 4: Detector measurements with gain factors applied (top to bottom) [1].

The latest UQ work as well as all other work from this project will be compiled together in the final release, BEAVRS v3.0, in the future.

3.2 BEAVRS Specifications - Version 2.0.1

BEAVRS v2.0 was released in September 6th, 2016. This release included a few improvements on the BEAVRS model as well as newly added data, for example, the inlet coolant temperature

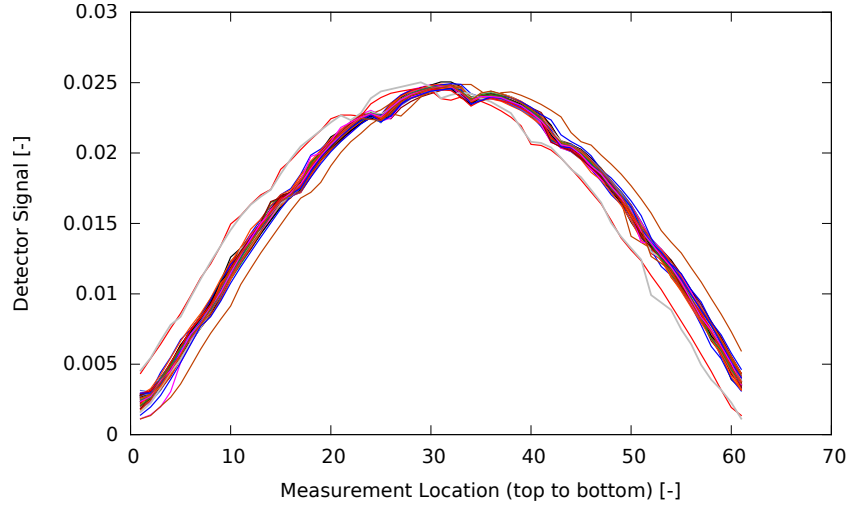


Figure 5: Detector measurements with zero points removed (top to bottom) [1].

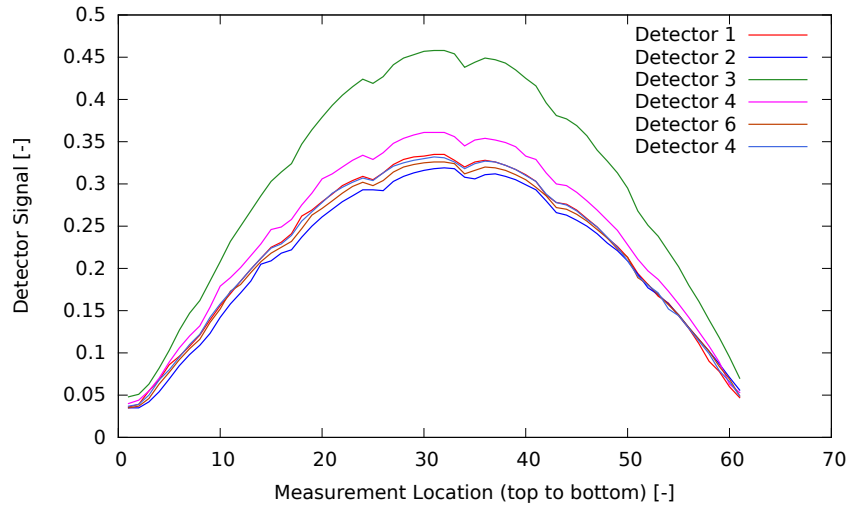


Figure 6: Detector measurements within assembly J10 (top to bottom) [1].

measurements. The densities of materials in the nozzle and support plate regions are changed to preserve the mass of steel and volume fraction of water when assuming a simplified geometry for the nozzles and support plate components. Also added was eighth-core symmetric radial data, or tilt-corrected radial data, which is generated by adjusting the detector measurements to correct the planar x-y tilt.

Version 2.0.1 of the BEAVRS specification contains several more updates according to feedback obtained from the benchmark users. For instance, CASL users reported some detector measured

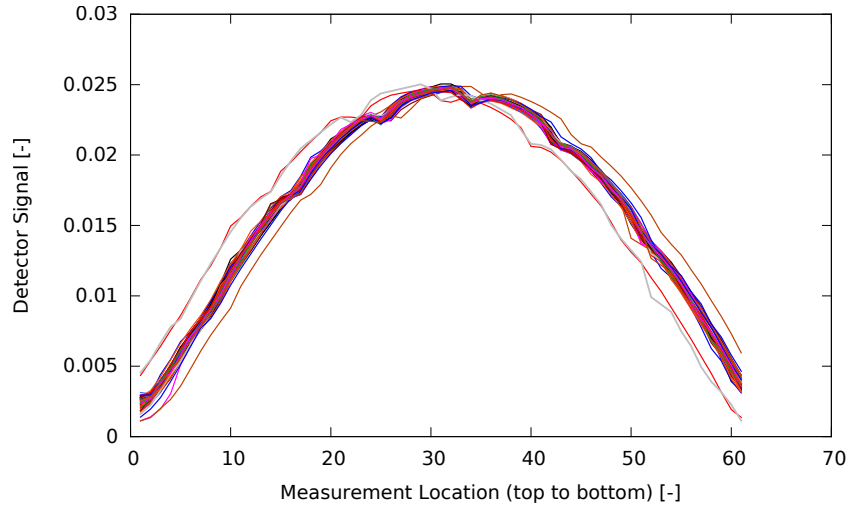


Figure 7: All detector signals before realignment [1].

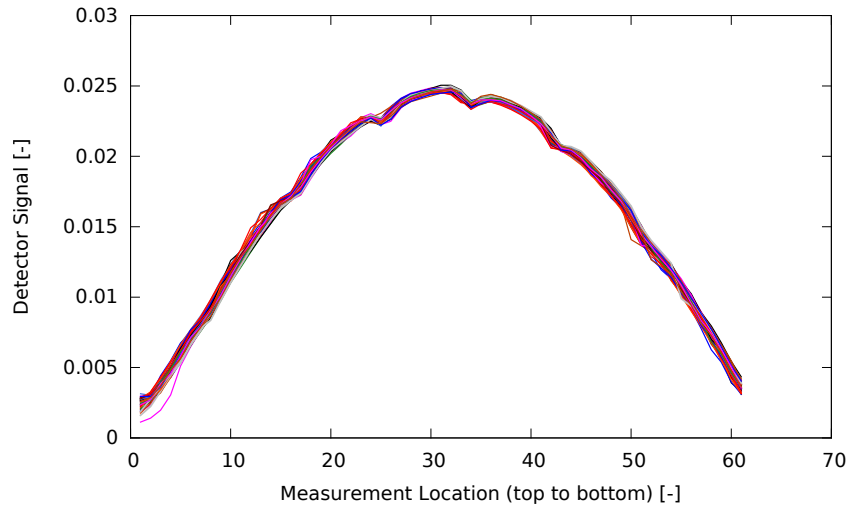


Figure 8: All detector signals after realignment [1].

axial signals are not correctly lined up when they were validating the VERA code. A more accurate and robust algorithm was then implemented to re-align the signals.

Detailed changes include:

- Slight shifts to all axial planes
- Added B4C control rod material and updated the control rod specification
- Updated the grid spacer specifications with more accurate heights

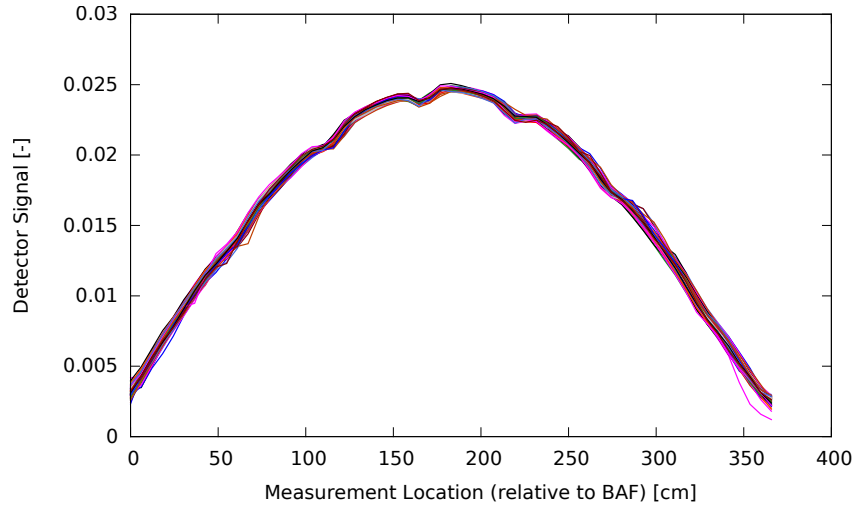


Figure 9: Comparison of all assemblies after spline [1].

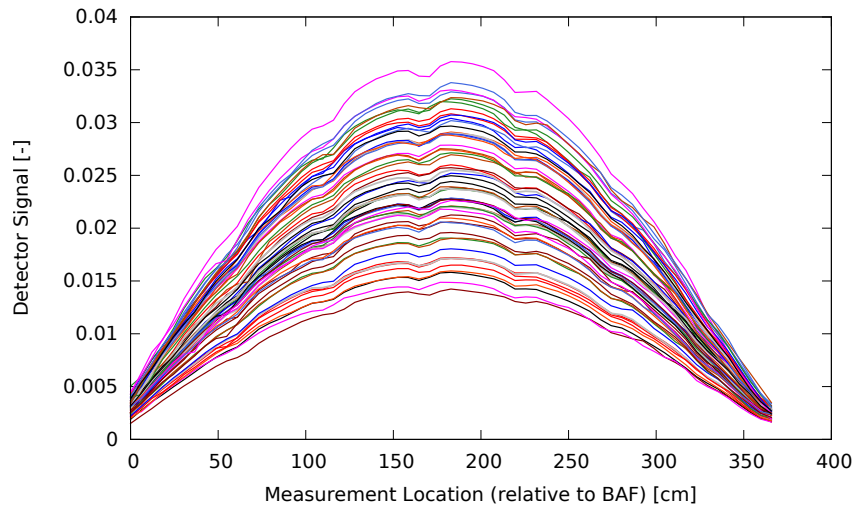


Figure 10: Final processed HZP measurement data [1].

- Updated the control rod 0% withdrawn axial location
- Updated the burnable poison insertion axial location
- Added upper plenum and end plugs to burnable poison rods
- Added upper plenum, end plugs, and spacer region to control rods
- Changed from air to helium for control rod and control rod gaps

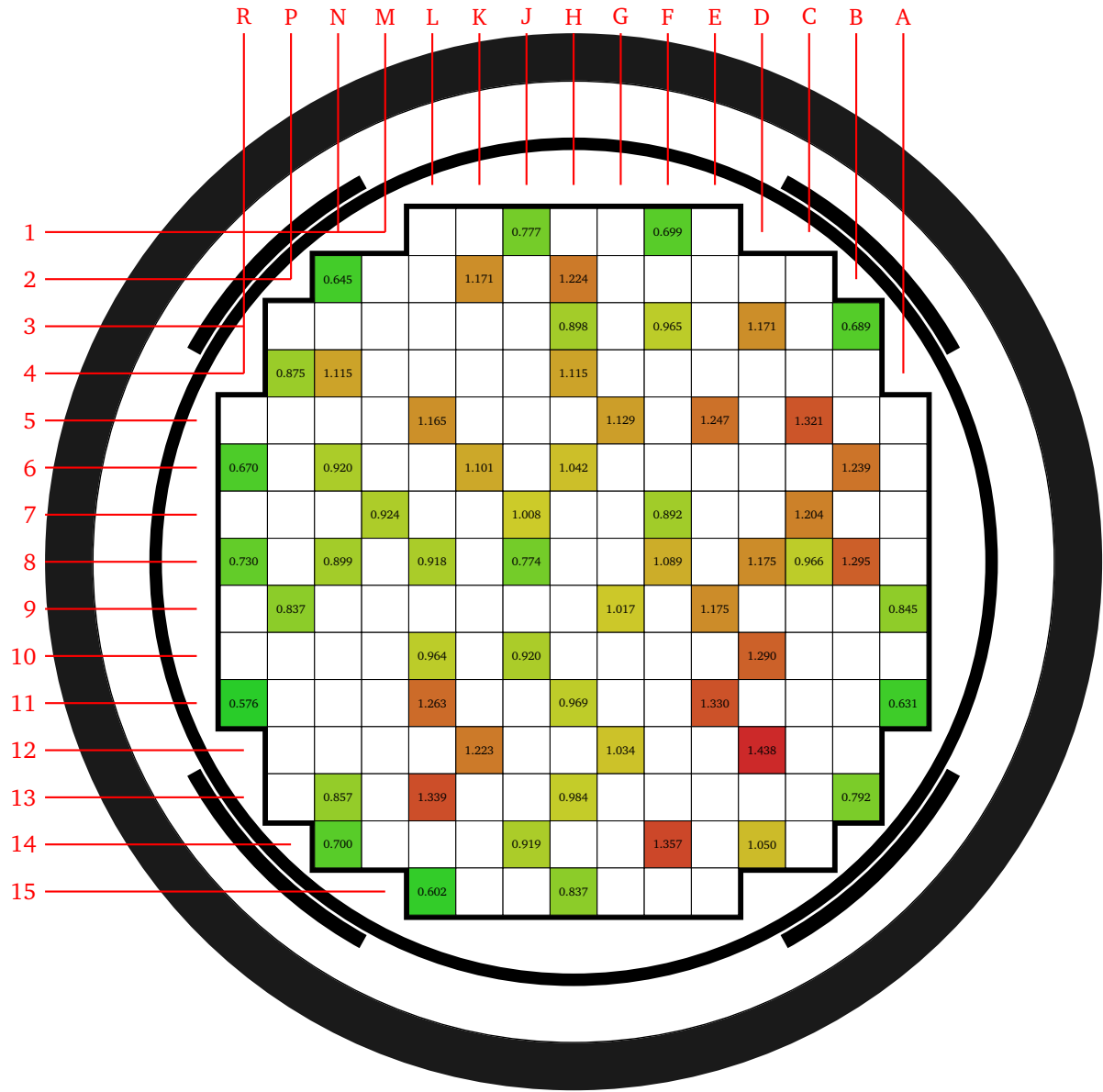


Figure 11: Radial detector measurements at HZP (axially integrated) [1].

- Updated neutron shield panel specification
- Updated RPV geometry and material specification
- Added the core liner
- Added water gap between fuel assemblies and baffle

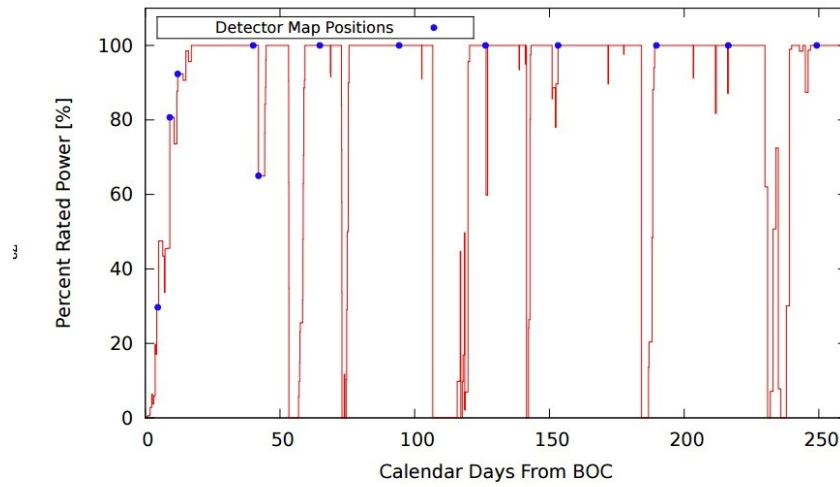


Figure 61: Power history of Cycle 2.

Figure 12: Cycle 2 power history in BEAVRS ver. 2.0.1

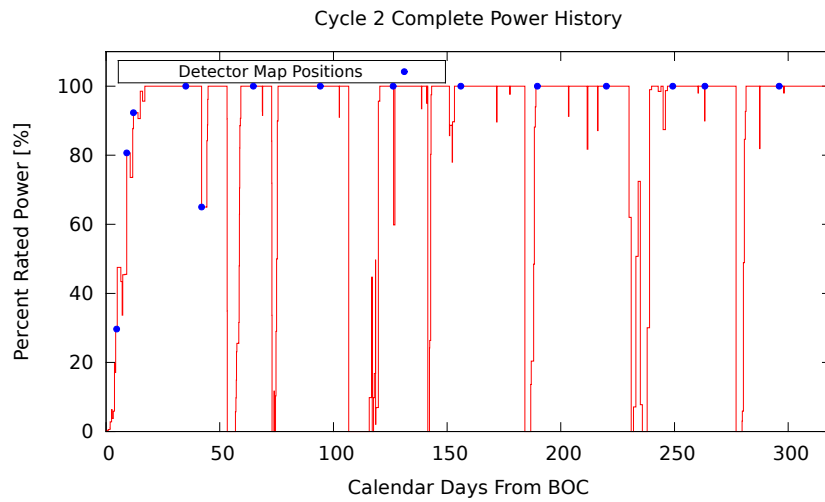


Figure 13: Updated cycle 2 power history in BEAVRS ver. 2.0.2

- Added new materials for nozzle and support plate structure to conserve mass and volume fractions
- Added inlet coolant temperature measurements
- Added tilt-corrected data

- Re-calculated all material compositions using latest isotopic natural abundances data, IUPAC 2013
- Corrected the critical Boron concentration from 1237 to 1273 ppm in Cycle 2
- Improved re-alignment algorithm in processing measurement data

The updates in the core structure are summarized in the figures below.

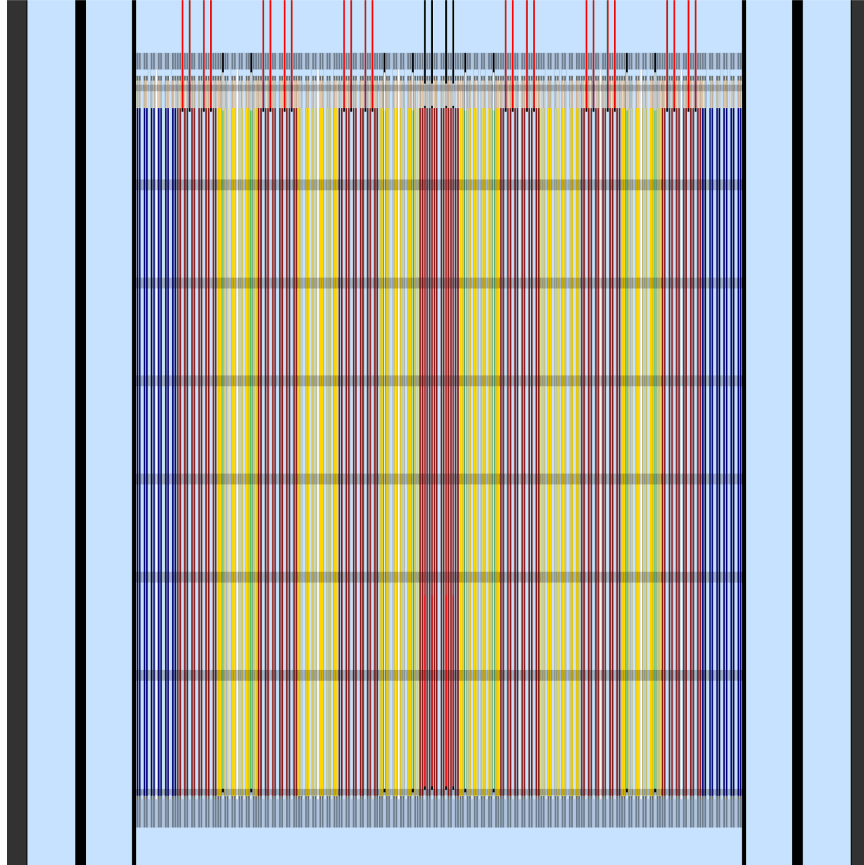


Figure 14: Scale view of row 8 axial cross section, with highlighted grid spacers and partial insertion of control rod bank D to the bite position

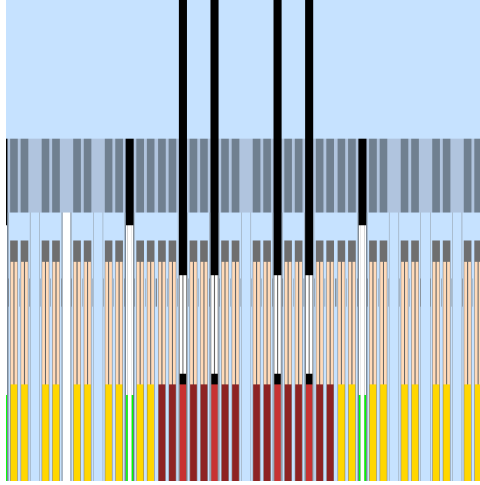


Figure 15: Axial scale view of the model near the top of the fuel rods in row 8, show pin plenums, approximated springs, end plugs, and structures. *Blue*: water; *orange*: helium; *black*: stainless steel; *dark gray*: Zircaloy; *dim gray*: Inconel; *white*: air; *slate gray*: nozzle / support plate stainless steel; *steel blue*: nozzle / support plate borated water; *green*: borosilicate glass; *yellow*: fuel.

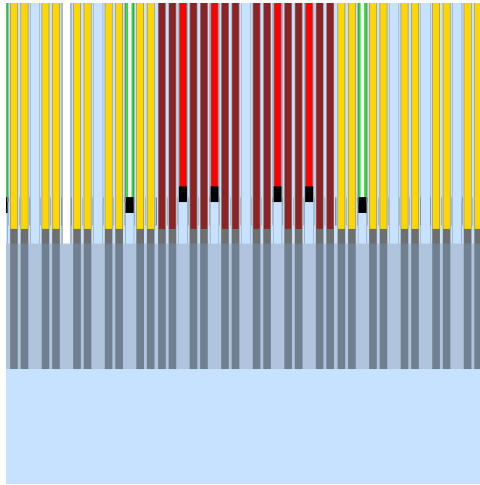


Figure 16: Axial scale view of the model near the bottom of the fuel rods in row 8, show pin plenums, approximated springs, end plugs, and structures. *Blue*: water; *orange*: helium; *black*: stainless steel; *dark gray*: Zircaloy; *dim gray*: Inconel; *white*: air; *slate gray*: nozzle / support plate stainless steel; *steel blue*: nozzle / support plate borated water; *green*: borosilicate glass; *yellow*: fuel.

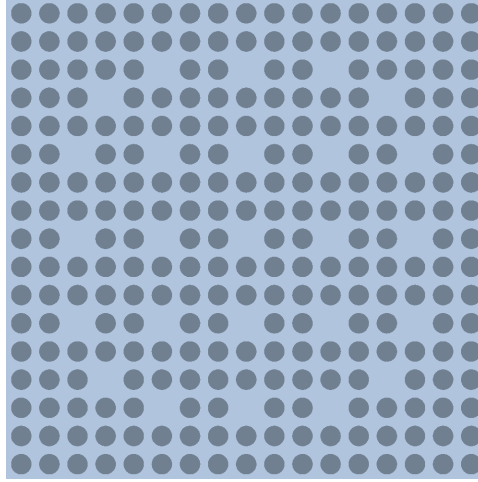


Figure 17: Radial picture of nozzles and support plate. The materials below and above the fuel rod regions are defined as "nozzle / support plate stainless steel" in color *slate gray* and "nozzle / support plate borated water" in color *steel blue* to conserve the mass the stainless steel and borated water in the nozzle and support plate region.

3.3 Benchmark Submission

This benchmark is proposed for inclusion in the international handbook, International Reactor Physics Experiment Evaluation Project (IRPhEP) of OECD/NEA. An initial version of the BEAVRS evaluation, identified as "BEAVRS-PWR-POWER-001 CRIT-REAC-COEFF-RRATE", was submitted to the IRPhEP Technical Review Group in September. We were invited to attend the IRPhEP Technical Review Meeting 2017 in October 24-26 in Washington, DC. The submitted BEAVRS benchmark was presented and reviewed. The benchmark will be included in the handbook once all comments are addressed.

The IRPhEP submission for BEAVRS was prepared based on existing BEAVRS documents. The previous specifications were re-formatted in the required Word format. Four types of measurements, critical configuration, reactivity effect measurements, reactivity coefficient measurements and reaction rate distribution measurements were presented. Also compiled was the uncertainty quantification analysis of all measurements. New computer codes simulations were incorporated as well, including OpenMC and CASMO/Simulate inputs and results.

4 Task: HZP Measurement Uncertainties

4.1 Assembly loading

As built assembly enrichment and loaded position are known for every assembly in the first 2 cycles of the BEAVRS benchmark. For modelling purposes it is simpler to consider only 3 different loading enrichment instead of 193 unique ones. The following tables summarize the average enrichment for each batch and standard deviation for the first two cycles.

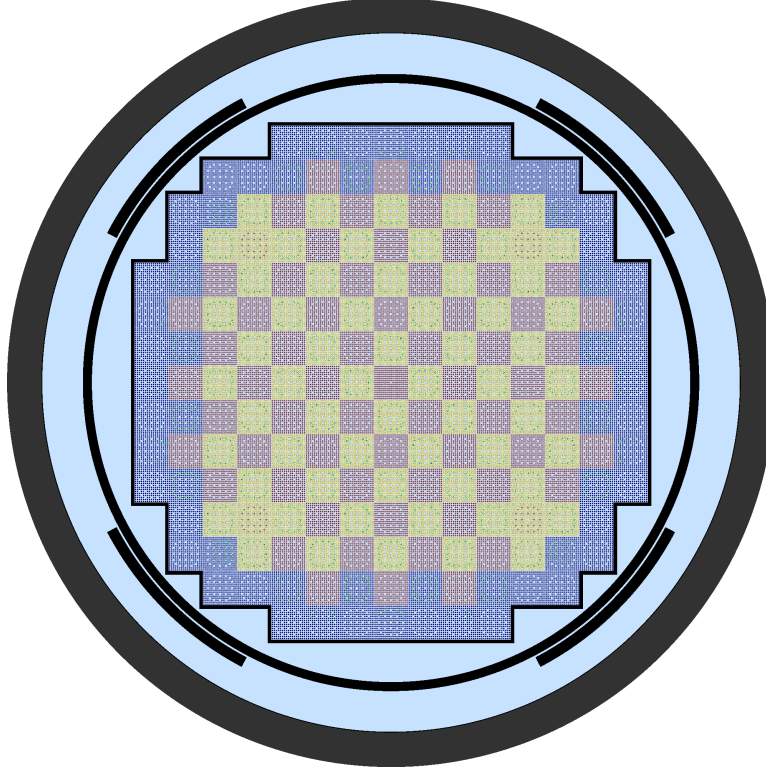


Figure 18: Cycle 1 core loading

	Fuel 1.6%	Fuel 2.4%	Fuel 3.1%
Number of assemblies	65	64	64
Average enrichment	1.610	2.400	3.102
Standard deviation	0.004	0.005	0.014

	Fuel 3.2%	Fuel 3.4%
Number of assemblies	48	16
Average enrichment	3.195	3.406
Standard deviation	0.005	0.003

4.2 Low Power Physics Tests

This report discusses the estimated uncertainties for the BEAVRS benchmark on the hot zero power (HZIP) physics tests. Since the BEAVRS measurements were performed a long time ago and very little data exists that can be used to evaluate the uncertainties, more recent data using similar techniques will be used to assess a best estimate of uncertainty to the BEAVRS data. This data set was provided by KHNP who started 4 OPR reactors with nominally identical cores. This data will be used to evaluate the uncertainties associated with measurements of critical boron

concentration, isothermal temperature coefficients and control rod worths. Obviously the OPR reactors are different to the BEAVRS model, but this modern data will still be used as the basis for measurement uncertainty since the measurement techniques are comparable. When possible, comparisons with other literature data points will be made to support this analysis. It should also be noted that many of the HZP measurements rely on the reactivity computer settings.

4.2.1 Critical Boron Concentration

In the OPR reactors, the critical boron concentration (CBC) was measured under three scenarios:

1. All rods out (ARO)
2. Shutdown bank B (SDB) in
3. Regulating group 1 (RG1) in

The process for measurement can be summarized as follows:

1. Borate the reactor coolant system (RCS) to withdraw all rods (for ARO case) except for regulating group 5 (RG5)
2. Withdraw RG5 partially and record position
3. Withdraw RG5 to upper limit and record reactivity ($\Delta RG5$)
4. Measure boron concentration (C_M)
5. Calculate CBC using inverse boron worth (IBW)

$$CBC = C_M + \frac{\Delta RG5}{IBW} \quad (1)$$

Looking at the measurement data from the 4 nominally identical OPR reactors, the measurement reproducibility is on the order of 8 ppm. However, this value does not include any uncertainties associated to the IBW obtained from simulation and uncertainties from the reactivity computer and its underlying parameters and assumptions. Thus the 8 ppm uncertainty can be attributed to the critical boron concentration measurement, but we must still account for the rod worth differential and the IBW. From the provided data, we can propagate the uncertainty of these parameters using the following relation:

$$\delta CBC = \sqrt{\left(\frac{\partial CBC}{\partial C_M} \delta C_M\right)^2 + \left(\frac{\partial CBC}{\partial \Delta RG5} \delta \Delta RG5\right)^2 + \left(\frac{\partial CBC}{\partial BW} \delta BW\right)^2} \quad (2)$$

which yields

$$\delta CBC = \sqrt{(\delta C_M)^2 + (BW \delta \Delta RG5)^2 + (\Delta RG5 \delta BW)^2} \quad (3)$$

The last term is quite small and can be neglected. We also know the boron worth (BW) from measurement. Assuming that the repeated measurements provide an estimate for C_M of 8 ppm, this leaves determining the reactivity change error associated with small movements of $RG5$. This error is associated with knowing the position of the rod and its associated reactivity effect for a

small change, and not the measurement error of the reactivity worth when fully withdrawn. Since the rod movement of RG5 is limited to approximately one tenth of its length, and assuming a measurement uncertainty of less than 5% (see control rod worth section), we can estimate the critical boron concentration uncertainty to be 22 ppm.

In the literature, a recent paper performing full core analysis of the Krsko plant reported critical boron measurement uncertainties of 20 ppm. This value is in line with the calculated value from the OPR data. Both values are also well below the common licensing criterion of 50 ppm for prediction from models.

Table 1: BEAVRS Cycle 1 Critical Boron Concentrations

Case	Critical Boron (ppm)	Std. Dev. (ppm)
ARO	975	22
D in	902	22
C,D in	810	22
A,B,C,D in	686	22
A,B,C,D,SE,SD,SC in	508	22

Table 2: BEAVRS Cycle 2 Critical Boron Concentrations

Case	Critical Boron (ppm)	Std. Dev. (ppm)
ARO	1405	22
C in	1237	22

4.2.2 Isothermal Temperature Coefficient

The isothermal temperature coefficient (ITC) test is performed using the following procedure:

1. Decrease temperature at a rate of 5 to 10 K per hour using the steam bypass control system
2. Maintain at 5 C below until data is recorded
3. Increase temperature at a rate of 5 to 10 K per hour using the steam bypass control system
4. Maintain at 5 C above until data is recorded
5. Return to temperature
6. Calculate RTC using changed in reactivity and reactor control system temperature for each change in temperature, and average results

In this procedure the ITC is thus measured 3 times and averaged. Only the average value is reported with its associated uncertainty. The OPR data contains 2 sets of measurements for all 4 reactors; one with all rods out (ARO) and one with shutdown bank B (SDB) fully inserted. For these cases, the ITC reported deviation is roughly 0.2 pcm/K.

The Palo Verde report [3] presents ITC measurements at cycle 1 for their three nominally identical cores. Using their data, the deviation is on the order of 0.6 pcm/K. In a recent paper,

the analysis performed on the Krsko NPP reports ITC measurement errors of 1.8 pcm/K. There thus exists very little agreement between the published literature and the measurements obtained from KHNP. The value from the Krsko plant seems unusually high and provides no supporting documentation. We thus have 11 measurements separated over three data sets. Weighing them by the number of data points in each set, we can approximate the ITC uncertainty to 0.3 pcm/K or 0.54 pcm/F.

Table 3: Palo Verde ITC data for Cycle 1 of all three units

Unit	ITC (pcm/F)
1	-4.40
2	-4.28
3	-3.79

Table 4: BEAVRS Cycle 1 ITC

Case	ITC (pcm/F)	Std. Dev. (pcm/F)
ARO	-1.75	0.54
D in	-2.75	0.54
C,D in	-8.01	0.54

Table 5: BEAVRS Cycle 2 ITC

Case	ITC (pcm/F)	Std. Dev. (pcm/F)
ARO	1.71	0.54

4.3 Control Rod Worth

At the time the BEAVRS reactor was started the most common way to measure control rod worths was with boron dilution followed by rod swap. Some of the early cores were done entirely by boron dilution, but to accelerate startup time, the rod swap technique became quite common once rod worth for one bank was known.

The boron dilution technique involves a decrease in boron concentration following a rod insertion in a sequence of small discrete steps. The reactivity computer print out is then used to calculate the incremental reactivity changes at each insertion. This method is very precised but very time consuming since sufficient wait time must be allowed during insertions for flux stabilization.

The rod swap technique requires that at least one bank be computed using the dilution technique to serve as a reference. The reference bank is often selected to be the highest worth bank since less rod displacement will be required. In this technique, the bank to be evaluated is compensated by the reference bank. The differential is then measured to calculate the effective reactivity with respect to the reference.

In the OPR reactors, the boron dilution technique is used for SDB and RG1, while rod swap is used for all others. Standard deviations were calculated for each rod bank and converted into relative error (standard deviation divided by average). Additionally, since the cores are not entirely

the same, predicted rod worths were used to include deviations in as-built fuel enrichment. However, correcting the experimentally measured values by the predicted ones, barely changed the uncertainty indicating that the measurement errors are much larger than the enrichment differences.

Figure 19 shows the trend line obtained from the data and the fitting function. Actual data points were omitted due to the proprietary nature of the data from KHNP. The red line indicates the plant acceptance criteria for deviations between the predictive models and measurements.

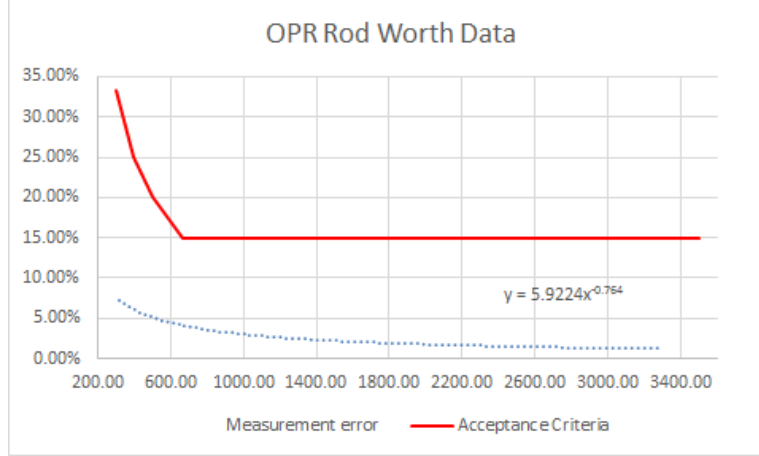


Figure 19: Trend line of rod worth measurement errors

Using this data with the BEAVRS reported rod worths provides an estimate of the measurement errors as shown in Table 6 and Table 7.

Table 6: BEAVRS Cycle 1 Rod Worths

Bank	Rod Worth (pcm)	Error (%)	Std. Dev. (pcm)
D	788	3.6	29
C	1203	2.6	32
B	1171	2.7	31
A	548	4.8	26
SE	461	5.5	25
SD	772	3.7	28
SC	1099	2.8	31

5 Task: Flux Map Uncertainties

This reactor contains 58 assemblies that can be accessed by in-core detectors, through the central guide tubes. Figure 20 shows these positions. There are six U-235 fission chambers with varying masses that are used to access the 58 positions. When measurements are being taken, multiple passes are performed to adequately measure all 58 assemblies. All detectors are passed through one

Table 7: BEAVRS Cycle 2 Rod Worths

Bank	Rod Worth (pcm)	Error (%)	Std. Dev. (pcm)
D	426	5.8	25
C	1014	3.0	30
B	716	3.9	28
A	420	5.9	25
SE	438	5.7	25
SD	305	7.5	23
SC	307	7.5	23
SB	781	3.6	29
SA	326	7.1	23

common assembly for signal normalization. The detectors are inserted from the bottom until they reach the top and pulled back to measure the axial fission rate over 61 axial locations.

5.1 Extracting detector signals and processing data

The raw measurement data for one detector in one pass includes 61 axial detector signals, background signal, gain factor of the detector and core power. The axial flux can be obtained from the detector signals by removing the background, adjusting the gain on the detector and dividing the power, as indicated in equation 4.

$$\phi_{ijk} = \frac{(D_{ijk} - B_{ij}) \times G_{ij}}{P} \quad (4)$$

where ϕ is the calculated flux, D is the detector signal, B is the background signal, G is the gain factor of the detector, P is the core power and i, j, k indicate the spatial position of a measurement (i, j for radial assembly position and k for axial location).

Radial assembly flux maps are obtained by integrating over all axial points, as indicated in equation 5.

$$\phi_{ij} = \sum_{k=1}^K \phi_{ijk} \quad (5)$$

However, the measurements are not usable as is and further post-processing is needed to filter the noisy data (such as missing data points, mis-alignments), which includes:

1. **Interpolation:** in some of the detector signals, zero points exist where the detector failed. These zero points are removed by performing a linear interpolation/extrapolation between/from the nearest two points.
2. **Realignment:** it is observed that not all of the signals are aligned with each other since the starting position of the recording can differ slightly. However, signals can be realigned according to grid depressions of the signal since grid positions are fixed.
3. **Spline fitting:** detector signals need to be put on an axial coordinate grid corresponding to points that range from the bottom to the top of the active fuel. A 2^{nd} order spline fit is used

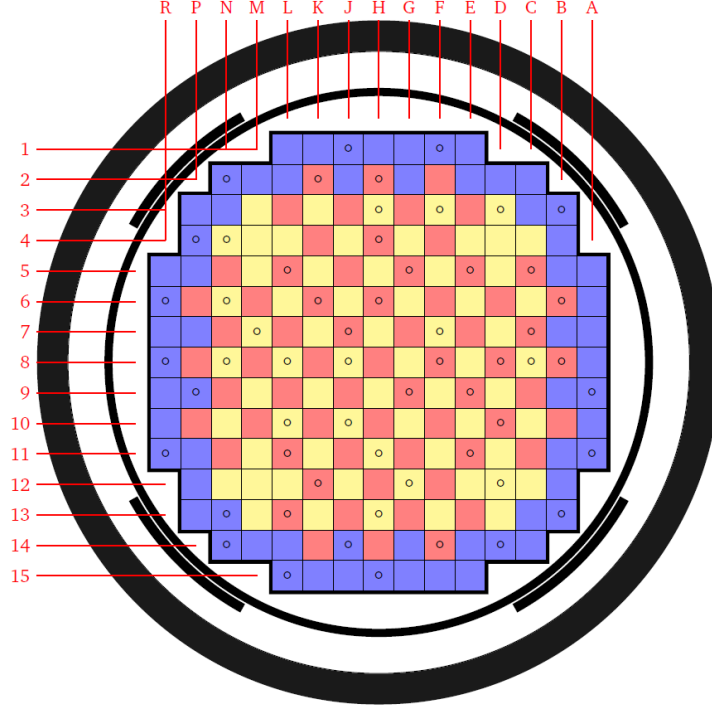


Figure 20: Instrument tube locations

to map from measured data axial locations to an axial map with data points exactly at the Top of Active Fuel (TAF) and Bottom of Active Fuel (BAF).

A Python script was developed to do the post-processing of the raw detector data. The final flux maps, including axial distributions and axially integrated radial maps of all 58 assemblies are generated by this script. For instance, the current hot zero power axially integrated power profile is shown in figure 21.

5.2 Estimating uncertainty of axial detector data

To evaluate uncertainties of the processed data, the uncertainties of the measured signals and those coming from post-processing are analyzed and combined.

The uncertainty of each axial detector data can be expressed as:

$$\begin{aligned}
 \left(\frac{\delta\phi}{\phi}\right)_m &= \sqrt{\frac{\left(\frac{\partial\phi}{\partial D} \cdot \delta_d\right)^2 + \left(\frac{\partial\phi}{\partial B} \cdot \delta_b\right)^2 + \left(\frac{\partial\phi}{\partial G} \cdot \delta_g\right)^2 + \left(\frac{\partial\phi}{\partial P} \cdot \delta_p\right)^2}{\phi^2}} \\
 &= \sqrt{\frac{\delta_d^2 + \delta_b^2}{(D - B)^2} + \left(\frac{\delta_g}{G}\right)^2 + \left(\frac{\delta_p}{P}\right)^2}
 \end{aligned} \tag{6}$$

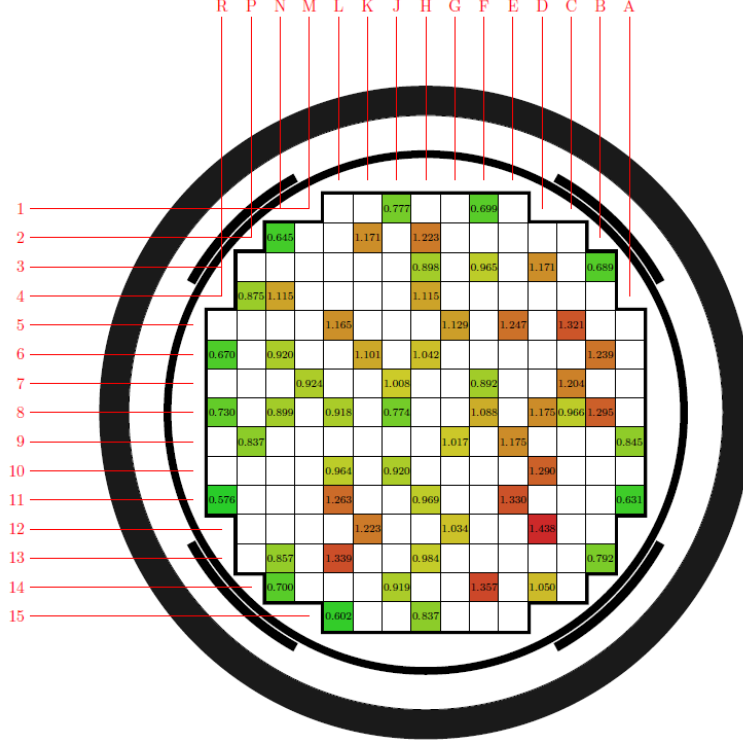


Figure 21: Hot zero power radial detector measurements (axially integrated).

$$\left(\frac{\delta\phi}{\phi}\right)_{ijk} = \sqrt{\left(\frac{\delta\phi}{\phi}\right)_m^2 + \left(\frac{\delta\phi}{\phi}\right)_{intp}^2 + \left(\frac{\delta\phi}{\phi}\right)_{align}^2 + \left(\frac{\delta\phi}{\phi}\right)_{spline}^2} \quad (7)$$

where δ_d , δ_b , δ_q , δ_p , the uncertainties of the detector signal, background, gain factor and core power respectively, contribute to the measurement uncertainty $\left(\frac{\delta\phi}{\phi}\right)_m$, while $\left(\frac{\delta\phi}{\phi}\right)_{intp}$, $\left(\frac{\delta\phi}{\phi}\right)_{align}$, and $\left(\frac{\delta\phi}{\phi}\right)_{spline}$, are the uncertainties introduced by interpolation, realignment and spline fitting, respectively.

5.2.1 Uncertainty from measurement

As shown in equation 6, the uncertainty of the axial flux has contributions from measuring the detector signal, the background, the gain and the core power. All types of measured data are gathered and statistical analysis is used to get the ranges and distributions of the measurements, as shown in figure 22 for cycle 1. It is found that:

1. The background values are very small, mostly provided by readings of 0 or 0.001. Accounting for the smallest division of the detector (0.001), uncertainty of the background can be

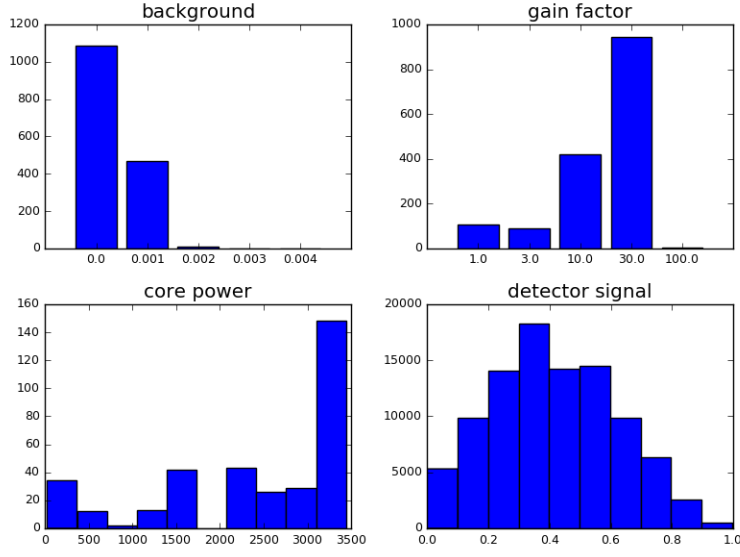


Figure 22: Hot zero power radial detector measurements (axially integrated).

estimated as half the smallest digit (0.0005) (i.e. $\delta_b = 0.0005$). It should be noted that background is subtracted from detector signal and the detector signal is generally greater than 0.1 (0.41 on average), which implies from equation 6 that both the effects of B and δ_b are less than 0.1% for most cases.

2. The gain factors are discrete values which are selected before each pass. As there are actually only a few well-spaced out values, it is believed that the detector error caused by using different gain factors is negligible compared to other sources, i.e. δ_g is 0.
3. For almost all of the measurements the core is operated with powers in the 100 or 1000 of MWth. As power is very important for normalization and values are reported with decimals, it can be safely assumed that its uncertainty (δ_p) is less than 1.0%, which also means that δ_p/P is under 0.1%.
4. Detector signal is the last remaining part. While the accuracy of detector is relative to specific measurements, literature indicates that for normal reactor applications that the uncertainty is on the order of 1% [4]. Therefore, it can be concluded qualitatively that the detector signal dominates the measurement uncertainty.

Now the question becomes how to evaluate the uncertainty of measured data more accurately. Quantitatively, measurement uncertainty can be evaluated by performing repeated measurements and calculating the variance. Even though the reactor core monitoring is mostly carried out only once for each assembly, there are indeed certain locations that receive multiple measurements for redundancy. For example in cycle 1 of BEAVRS, there are about 180 cases each that have 2 or even 3 repeated measurements by the same detector in the same assembly. These multiple measurements are performed with the reactor core under almost the same conditions, such as

control rod positions, boron concentration and coolant status. Core power as well as other measured parameters such as background, gain factor of detector may be slightly different but they will be accounted for in calculating axial flux. So the multiple measurements can be regarded as repeated data and their uncertainty represents the accuracy of the measured data. Figure 23 is an example of multiple measurements which show the independent axial signals, mean and relative sample standard deviation (RSTD) of every axial measurement point.

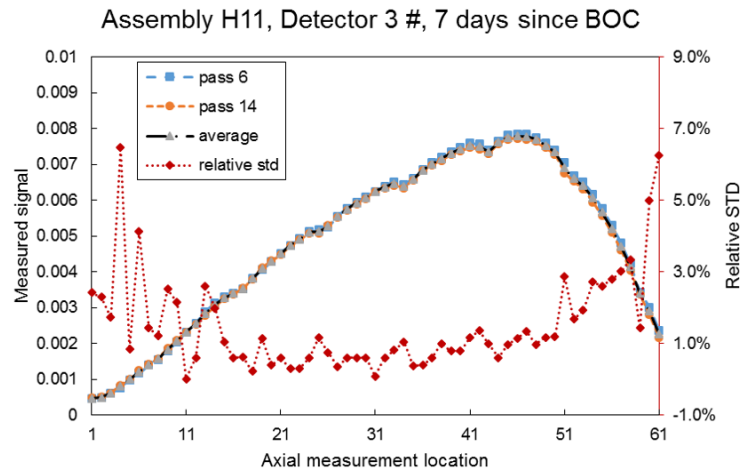


Figure 23: Multiple measurements example of H11 assembly

Since there are mostly only 2 replicated measurements for each data point, the variance estimate is not likely to be accurate, thus gathering the data over multiple cases through the whole cycle is necessary to obtain a reliable uncertainty estimate.

Here we assume the measurement uncertainty is related only to the detector signal amplitude. That means the multiple measurement cases with similar mean value have similar uncertainty, therefore we can group the data points by amplitude and estimate an average uncertainty for each group. Specifically, the relative sample standard deviations are gathered and used to calculate the uncertainties of all groups. To make the estimate more conservative, a 95% confidence value is adopted as resulting uncertainty. In calculation, all the data are sorted and the value at the 95% position is selected. Note that the data points in one replicated measurement case follow the same Gaussian distribution. For each double repeated measurements case, it can be demonstrated that the relative sample standard deviation follows a distribution given by equation 8 and the 95% confidence value is actually equal to 2 times of true standard deviation i.e. 2σ .

$$f(s) = \frac{4}{\sqrt{2\pi}\sigma} e^{-\frac{s^2}{\sigma^2}} \quad (8)$$

There are 11145 multiple measurements in cycle 1 and they are divided into 30 groups by signal amplitude with equal number of data points in each group. The 95% confidence values are calculated for each group one by one, thus a table of signal ranges and measurement uncertainties is obtained, as shown in figure 24 and 25.

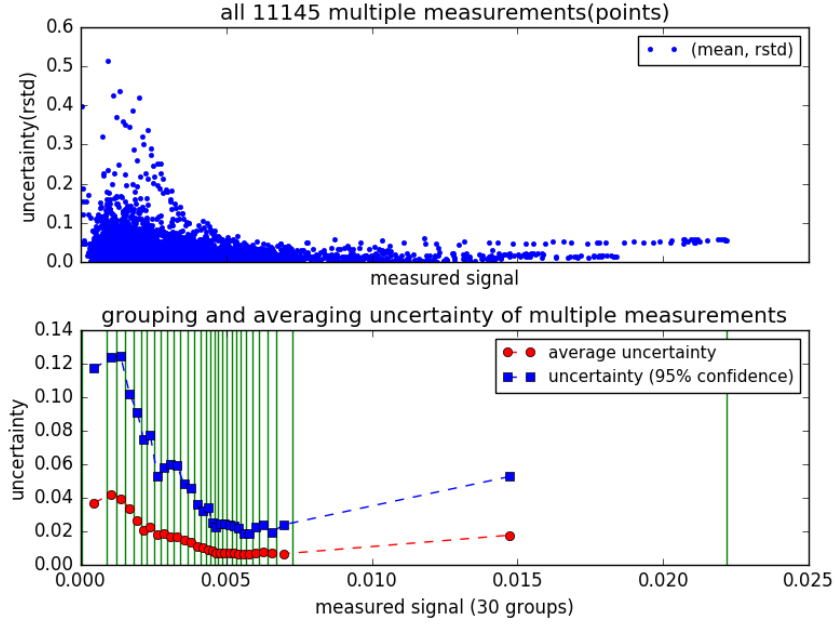


Figure 24: Axial signal uncertainty from all multiple measurements in cycle 1

As can be found in figure 24, the axial measurement uncertainty is generally around 2 to 12% and is dependent on signal amplitude, i.e. the greater the signal, the smaller the uncertainty. This look-up table can be applied to all detector signals to get the measurement uncertainties.

5.2.2 Uncertainty from Processing

To analyze the uncertainty caused by data post-processing such as interpolation, realignment and spline fitting, a method of observing the variation before and after the processing is used.

Interpolation

The interpolation is performed when one data point is unrecorded. To estimate the error of interpolated data, we just eliminate one-by-one points that were properly recorded and then compare the axial relative error with the recorded data. Figure 26 displays all the interpolation errors and their averages along with axial locations.

Interpolation error is dependent of location and it is greater near endpoints or grids. Here again we use the 95% confidence strategy for each location. The 95% confidence values are collected as a look-up table to estimate the uncertainty of every interpolation performed during processing.

Axial re-alignment

Re-alignment of axial signals is carried out such that axial signals are provided over the active fuel length. But it should be noted that only the positions of the axial data are moved up or down in re-alignment. The data points themselves are kept unchanged and are believed accurate, except for

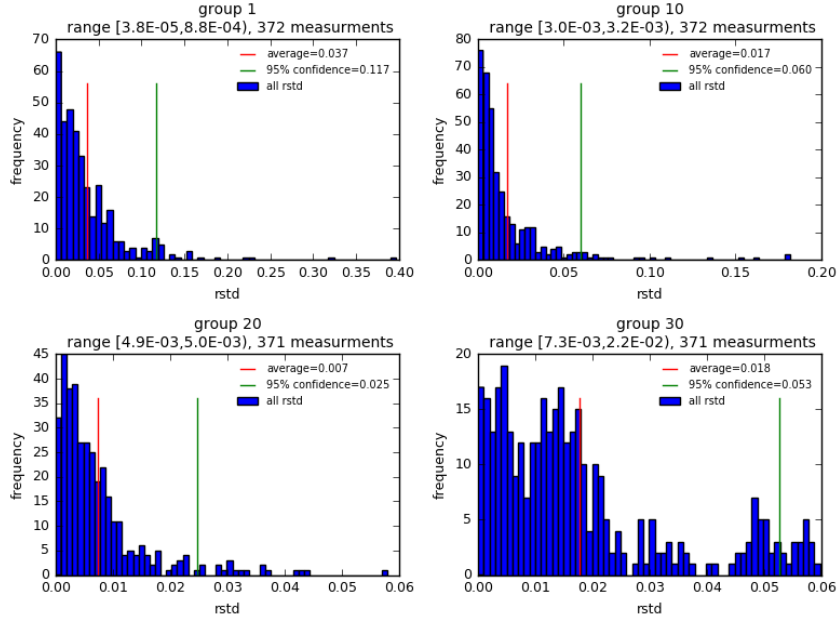


Figure 25: Average and 95% confidence values for each multiple measurements group

the lost edge points which are calculated using extrapolation. Therefore, the error of re-alignment essentially comes from extrapolation of end points. A similar method to interpolation is used to estimate the error of these extrapolations, i.e. comparing normal measured data with extrapolated data. Axial measurements are re-aligned by 1 or 2 axial positions and in rare situations 3. Figure 27 shows an example of the effect of grid re-alignment.

Figure 28 shows the error of axial re-alignment i.e. the relative error caused by extrapolation in 6 grid positions and the means of absolute errors. Again, these data will be used as a look-up table to estimate error of axial points which are extrapolated when re-alignment is performed.

Spline fitting

Spline fitting is used to put the detector signals on an axial coordinate grid corresponding to points that range from the bottom to the top of the active fuel. The error introduced in spline fitting is found negligible because the 2nd order spline fit is quite accurate and the errors are less than 10^{-10} .

$$\left(\frac{\delta \phi}{\phi} \right)_{spline} \approx 0 \quad (9)$$

5.2.3 Combining uncertainties

Finally, the uncertainties of all axial signals can be quantified by calculating every independent uncertainty using the look-up tables and combining them together by the uncertainty equation i.e. in the following steps:

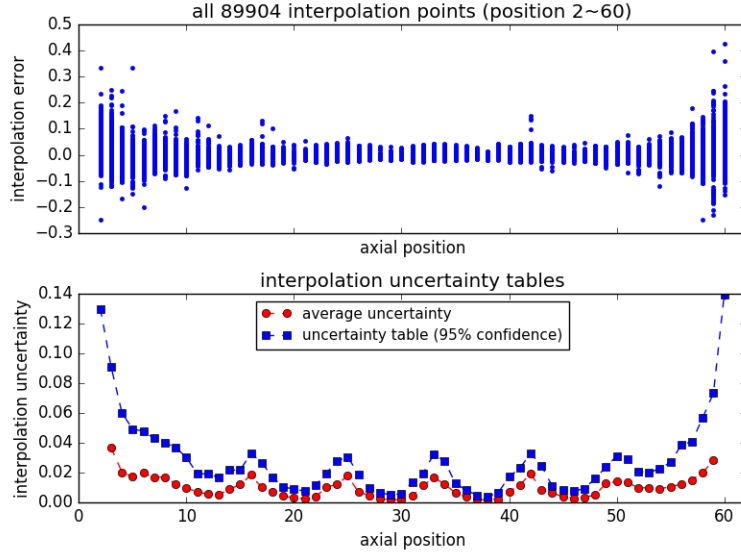


Figure 26: Axial interpolation error

1. The measurement uncertainty $\left(\frac{\delta\phi}{\phi}\right)_m$ is looked up according to the amplitude of calculated axial data.
2. If this data point is interpolated, an interpolation error $\left(\frac{\delta\phi}{\phi}\right)_{intp}$ is looked up according to its location.
3. If this data point is realigned, an extrapolation error $\left(\frac{\delta\phi}{\phi}\right)_{align}$ is looked up according to its location.
4. The square root of sum of squares of all uncertainties is calculated as resulting uncertainty $\left(\frac{\delta\phi}{\phi}\right)_{ijk}$ for each data point.

Figure 29 shows an example of combined axial signal uncertainty for one assembly in cycle 1.

Since most of the measured data points are not affected by interpolation or realignment, the uncertainty of axial detector data is generally between 2 to 12%.

5.3 Uncertainty of axially integrated radial data

Uncertainty of radial assembly data which is axially integrated can be evaluated by accounting for the errors of all axial points in one assembly, as indicated in equation 10.

$$\left(\frac{\delta\phi}{\phi}\right)_{ij} = \frac{\sqrt{\sum_{k=1}^K (\delta\phi)_{ijk}^2}}{\phi_{ij}} \quad (10)$$

where i, j represent radial assembly position and k for axial location.

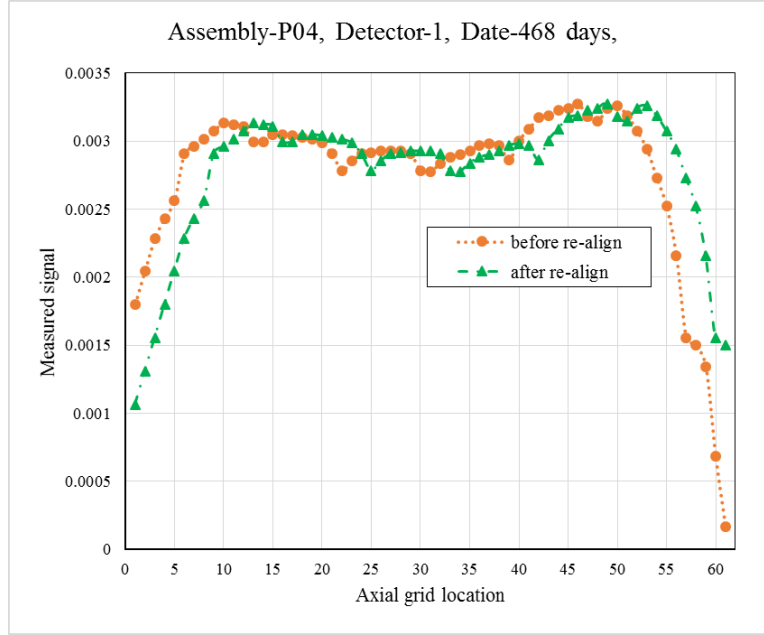


Figure 27: Example of large re-alignment error

For example, in assembly L05, the axial uncertainty shown in figure 29 can be substituted into the equation and an assembly uncertainty of 0.43% is obtained.

Table 8 gives full core radial detector measurements and calculated uncertainties on day 66 in cycle 1, while figure 30 gives the distribution of all assembly uncertainties in cycle 1. It can be seen the average uncertainty of radial measurements is around 0.6% and 95% confidence value is 1.2%.

The multiple measurements can also be used to quantify uncertainty of radial data as a second means. Similar to the treatment with the axial multiple signals, all multiple measurement cases of radial data are collected and statistics is made to estimate uncertainty of radial data, as shown in figure 31. There are 186 cases of multiple measurements in cycle 1. The average relative error is about 0.5% while the 95% confidence value is 1.5%.

It should be noted the two ways to estimate radial uncertainty are somewhat independent with each other so verification can be performed between the two. The two approaches are quite consistent with each other, as shown in figure 30 and figure 31. Therefore, it can be concluded the methodology to quantify the flux map uncertainties is basically correct and valid.

6 Task: HFP Boron Letdown Uncertainties

The measured boron letdown curves during Cycle 1 and Cycle 2 operation are given in the BEAVRS benchmark, as shown in Figure 32.

From literature review, the typical acceptance criteria from regulators is that codes must predict critical boron at ± 50 ppm. Recent measurements from low power physics tests of OPR reactors indicate measurement uncertainties below 1% or about 10 ppm.

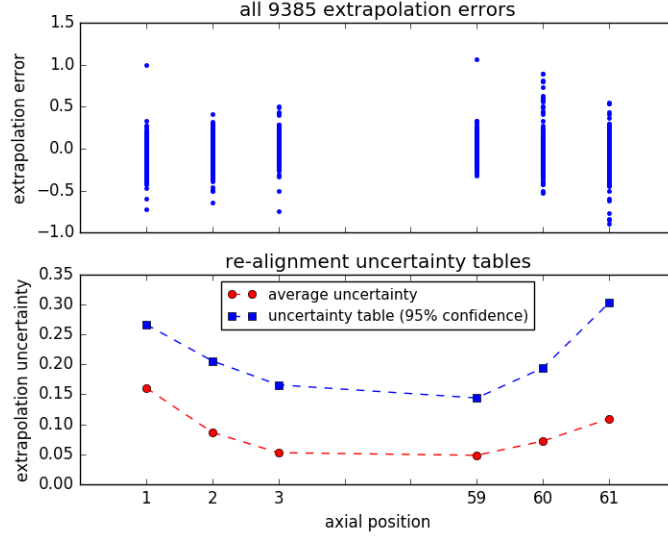


Figure 28: Axial re-alignment error

One way to evaluate boron measurement uncertainties is to estimate the accuracy of the boron meter (BM) used in the measurements. It is found in the original data files that multiple boron readings exist in some cases. Specifically, boron data is recorded for each detector pass thus the count of boron readings is equal to the number of detector passes, as shown in Table 9. Since the measuring passes are carried out on the same day with control rods at the same positions, i.e. under almost the same core conditions, these multiple boron readings can be regarded as repeated independent measurements. Therefore, the standard deviation of the boron data should reveal to some extent the uncertainty of boron measurements.

The accuracy of boron instrument can be calculated from the statistical properties of the repeated measurements, as shown in the following equations.

$$X = \frac{1}{N} \sum_{i=1}^N x_i \quad (11)$$

$$\Delta X = \frac{1}{N} \sqrt{\sum_{i=1}^N \Delta x_i^2} = \frac{1}{\sqrt{N}} \Delta x \quad (12)$$

$$\Delta x = \sqrt{N} \Delta X \approx \sqrt{N} \cdot STD \quad (13)$$

where x is the boron measurement, X is the average of multiple measurements, N is the number of independent measurements, Δx is the accuracy of boron meter (BM) and ΔX is the uncertainty of multiple measurements which can be represented by the standard deviation (STD).

Table 10 gives estimations of BM accuracies for all the 19 multiple measurements cases. The average of all instances is 24.3 ppm. It should be noted that even though the boron meter accuracy is fluctuating due to the insufficient repeated measurements for every single case, the average of

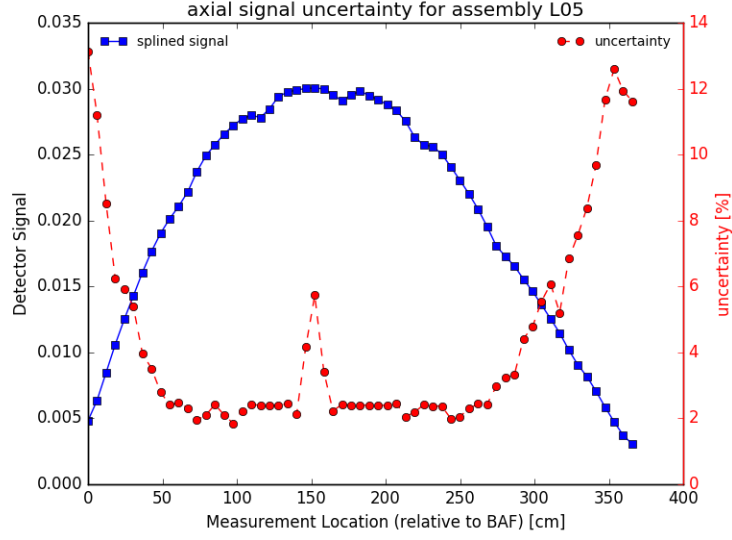


Figure 29: Combined uncertainty of axial detector measurement - assembly L05 on day 54

all these instances provides a reasonable estimate because the number of total cases (273) is quite large. Thus, the preliminary estimation of boron measurement uncertainty is around 25 ppm from the current approach.

An alternative way to quantifying the boron letdown uncertainty is to compare measurements with the code calculation. Here, CASMO-5/Simulate-3 codes are used to model the BEAVRS benchmark and boron letdown values are calculated along with the benchmark burnup points. Table 11 presents the measured and calculated values and their discrepancies in Cycle 1. It can be seen that the average discrepancy is 25.2 ppm, which has good agreement with the above boron measurement uncertainty estimated from multiple measurements. Therefore, the boron uncertainty of 25 ppm seems appropriate.

Table 8: Full core radial detector measurements and axially combined uncertainties on day 66

K12	$1.259 \pm 0.39\%$		D10	$1.271 \pm 0.39\%$
D12	$1.209 \pm 0.39\%$		H3	$0.941 \pm 0.55\%$
H2	$1.199 \pm 0.41\%$		L10	$1.047 \pm 0.48\%$
H6	$1.254 \pm 0.40\%$		L13	$1.256 \pm 0.40\%$
F7	$1.031 \pm 0.49\%$		F1	$0.643 \pm 0.91\%$
G9	$1.234 \pm 0.41\%$		F3	$0.950 \pm 0.55\%$
F8	$1.253 \pm 0.40\%$		E11	$1.328 \pm 0.38\%$
A11	$0.544 \pm 1.01\%$		R11	$0.519 \pm 1.06\%$
E9	$1.290 \pm 0.38\%$		G12	$1.046 \pm 0.48\%$
M7	$1.018 \pm 0.49\%$		N8	$0.939 \pm 0.55\%$
L15	$0.531 \pm 1.02\%$		N6	$0.936 \pm 0.55\%$
J10	$1.048 \pm 0.48\%$		N2	$0.569 \pm 0.95\%$
K2	$1.145 \pm 0.42\%$		H13	$0.973 \pm 0.52\%$
K6	$1.294 \pm 0.38\%$		L8	$1.032 \pm 0.50\%$
H11	$1.226 \pm 1.06\%$		L5	$1.284 \pm 0.38\%$
E5	$1.305 \pm 0.38\%$		H15	$0.725 \pm 0.81\%$
D8	$1.258 \pm 0.40\%$		C8	$0.954 \pm 0.54\%$
B3	$0.576 \pm 0.94\%$		B6	$1.159 \pm 0.42\%$
B8	$1.208 \pm 0.41\%$		H4	$1.248 \pm 0.40\%$
C5	$1.250 \pm 0.40\%$		F14	$1.217 \pm 0.41\%$
C7	$1.212 \pm 0.41\%$		N13	$0.708 \pm 0.83\%$
A9	$0.738 \pm 0.79\%$		N14	$0.585 \pm 0.93\%$
R6	$0.630 \pm 0.92\%$		R8	$0.690 \pm 0.87\%$
J8	$0.936 \pm 0.55\%$		P4	$0.768 \pm 0.78\%$
J1	$0.712 \pm 0.83\%$		P9	$0.798 \pm 0.71\%$
J7	$1.219 \pm 0.41\%$		B13	$0.604 \pm 0.92\%$

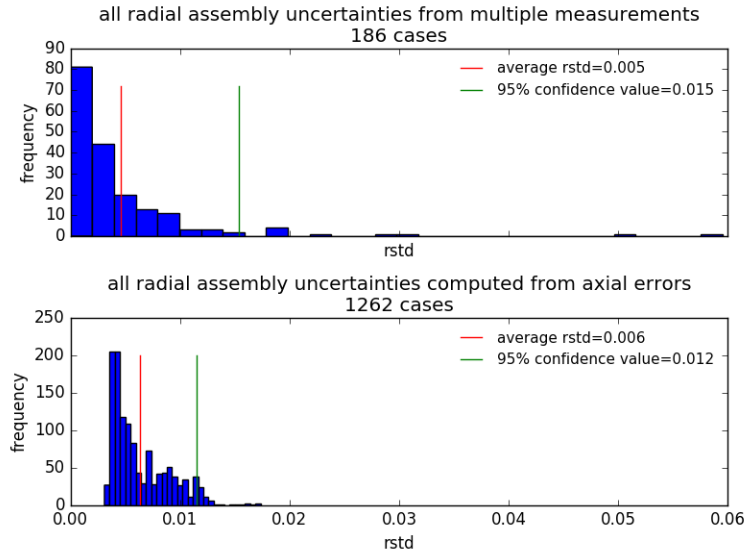


Figure 30: Distribution of radial data uncertainty calculated from axial uncertainties in cycle 1

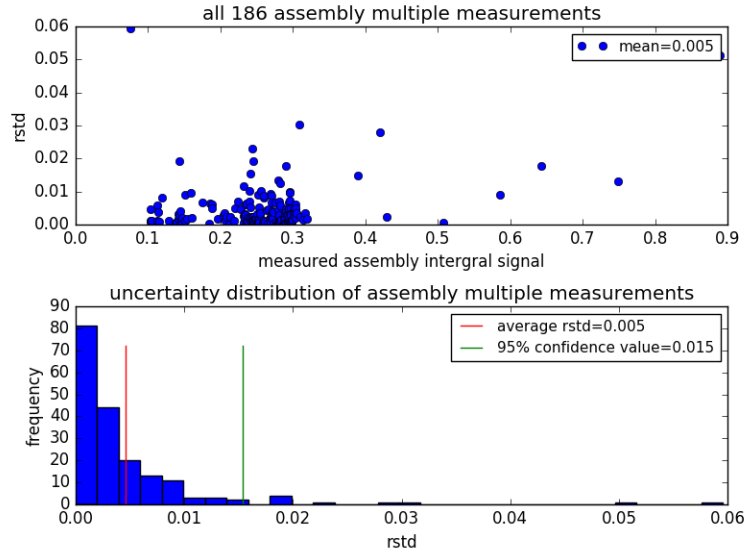


Figure 31: Distribution of radial data uncertainty using multiple measurements in cycle 1

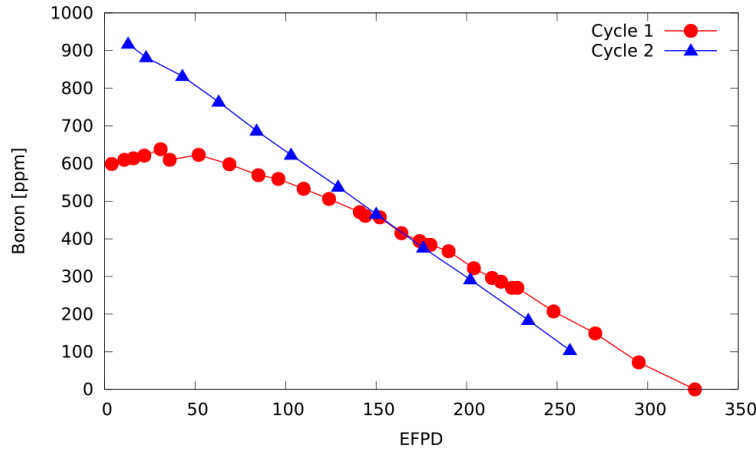


Figure 32: Boron letdown curves for two cycles of operation

Table 9: Repeated Boron measurements in cycle 1

days	# of passes	Boron measurements (ppm)
0	20	973, 964, 964, 991, 969, 972, 962, 977, 984, 984, 981, 986, 956, 986, 986, 991, 989, 973, 933, 984
7	14	881, 880, 881, 875, 878, 873, 877, 877, 881, 877, 875, 873, 870, 877
18	14	792, 798, 795, 805, 809, 802, 800, 802, 798, 806, 800, 803, 809, 800
54	13	744, 761, 764, 753, 753, 752, 756, 769, 750, 745, 752, 762, 750
62	17	705, 708, 698, 695, 703, 700, 700, 691, 711, 708, 705, 703, 706, 706, 703, 709, 702
66	25	700, 708, 695, 702, 706, 700, 708, 697, 706, 692, 698, 697, 706, 703, 694, 695, 697, 706, 697, 703, 695, 695, 697, 697, 706
81	13	645, 645, 658, 648, 658, 647, 661, 645, 662, 644, 652, 658, 655
82	13	642, 648, 645, 645, 647, 648, 648, 647, 645, 647, 647, 641, 616
88	13	630, 631, 639, 634, 630, 634, 634, 630, 631, 630, 631, 631, 636
92	13	608, 606, 627, 631, 634, 633, 631, 622, 634, 617, 623, 616, 612
161	16	695, 692, 703, 702, 698, 700, 705, 700, 702, 702, 702, 700, 691, 689, 700, 695
169	14	631, 630, 628, 633, 630, 627, 630, 630, 628, 628, 630, 633, 630, 630
187	14	622, 620, 619, 619, 617, 641, 620, 620, 619, 619, 622, 622, 620, 614
218	13	575, 577, 578, 580, 577, 580, 580, 578, 578, 575, 581, 580, 578
251	13	520, 514, 520, 516, 523, 516, 514, 514, 516, 519, 514, 520, 514
323	11	531, 530, 525, 527, 522, 523, 530, 523, 527, 528, 522
339	12	437, 436, 436, 433, 439, 436, 442, 434, 441, 436, 439, 434
368	14	439, 430, 423, 377, 362, 359, 369, 366, 364, 373, 373, 375, 380, 375
403	11	306, 305, 305, 305, 309, 305, 308, 305, 308, 305, 306

Table 10: Calculating boron meter accuracy from multiple measurements

days	N	Standard deviation (ppm)	BM accuracy (ppm)
0	20	14.4	64.4
7	14	3.4	12.6
18	14	4.9	18.3
54	13	7.4	26.7
62	17	5.2	21.3
66	25	4.9	24.7
81	13	6.8	24.4
82	13	8.6	30.9
88	13	2.8	10.1
92	13	10.0	36.1
161	16	4.7	18.9
169	14	1.7	6.5
187	14	6.1	22.9
218	13	1.9	6.9
251	13	3.1	11.2
323	11	3.4	11.2
339	12	2.8	9.7
368	14	26.6	99.4
403	11	1.5	5.0
average			24.3

Table 11: Comparison of measured boron level with CASMO/Simulate calculated boron in Cycle 1

Burnup (MWd/kg)	Measured Boron (ppm)	Calculated Boron (ppm)	Absolute error (ppm)
0	975.25	958	17.3
0.04	801.36	784	17.4
0.15	754.69	712	42.7
0.27	703.12	670	33.1
0.33	700	666	34.0
0.64	652.15	626	26.2
0.67	643.54	626	17.5
0.88	632.38	611	21.4
1.02	622.62	598	24.6
1.3	698.5	650	48.5
1.51	629.86	602	27.9
2.16	621	588	33.0
3.3	578.23	562	16.2
4.61	516.92	503	13.9
6.01	526.18	493	33.2
6.49	436.92	418	18.9
7.51	383.21	365	18.2
8.7	306.09	298	8.1
9.8	259	231	28.0
11.08	180	153	27.0
12.34	120	74	46.0
12.92	50	64	14.0
13.6	35	48	13.0
average			25.2

7 Task: Time Series Analysis

The BEAVRS Year 2 Report [5] develops the framework for quantifying time-dependent uncertainty using the linear model as well as the CASMO/Simulate model, and the Year 3 Report [6] expands on this work to attribute overall uncertainty values to such models. This uncertainty is composed of error that arises from assuming a linear tilt-correction to BEAVRS data, which is discussed in Section 7.1. In addition, Sections 7.2 and 7.3 look at the second source of time-dependent uncertainty - detection uncertainty - which is induced from fitting a simulation model to assembly-level BEAVRS data. In the case of Section 7.2, the simulation model under consideration is the CASMO/Simulate model, while Section 7.3 examines the validity of using the linear regression model instead. Finally, Section 7.4 summarizes the findings from this task.

7.1 Uncertainty in Using Tilt-Corrected BEAVRS Data

The BEAVRS Year 1 Report [7] introduces the need for a tilt corrected BEAVRS data set. The rationale for this can be understood by looking at Figure 21, which plots the original BEAVRS measured data at HZP. It can be seen that the reaction rate data is not eighth-core symmetric despite a symmetric core loading pattern, with the exception of instrumentation tubes and as-built masses. This tilt was first described in detail by Sykora in his Master’s thesis, where he devises an algorithm to remove this tilt to yield eighth-core symmetric reaction rates [8]. The magnitude of this tilt is not fixed, however, and slowly decreases over burnup. Eliminating the tilt is essential in the process of fitting the CASMO/Simulate model to BEAVRS data, which is explored in Section 7.2. This is because simulation tools will produce symmetric results given a symmetric core loading pattern and cannot arbitrarily impose a tilt to mimic that observed in the BEAVRS data. Thus, tilt correction allows for a more accurate comparison between operational and simulated data, but artificially masks an un-modeled phenomenon within the operating data. The exact amount of added uncertainty from assuming a linear tilt correction is the focus of this section. A linear tilt performed best in terms of modeling the tilt observed within operating data. Section 7.1.1 explores the process of identifying the best x-y plane to fit BEAVRS data, while Section 7.1.3 devises a comprehensive method for simulating the tilt observed in the BEAVRS data. Finally, Section 7.1.4 uses the radial map that results from this revised approach as a basis for evaluating tilt-correction uncertainty.

7.1.1 Generation of Tilt-Corrected Data

Asymmetry in the calculated measured fission rates are corrected by assuming that deviations follow a purely linear tilt. This linear tilt is determined by perturbing x-y tilt coefficients and finding the plane that best fits measured data. This plane minimizes RMS error between symmetric detector fission rates relative to the plane, and octant-core symmetric “tilt-corrected data” can be generated by subtracting the tilt of this plane from measured data. All modeled data will be compared to this tilt-corrected data. Figure 21 shows the original BEAVRS measured data at HZP, while Figure 33 illustrates what this data looks like after tilt-correction. The source of this behavior is not known exactly, but the leading hypothesis is that uneven water gaps during core loading is causing this irregularity [7]. Typically, core loading is performed in a crescent pattern starting from one corner and working around to the opposite corner. As a result, bundles could have leaned against the baffle during insertion, causing uneven gaps to occur in opposite corners of the core. As the reactor heats up, symmetric water gaps are restored through fuel swelling under irradiation and the

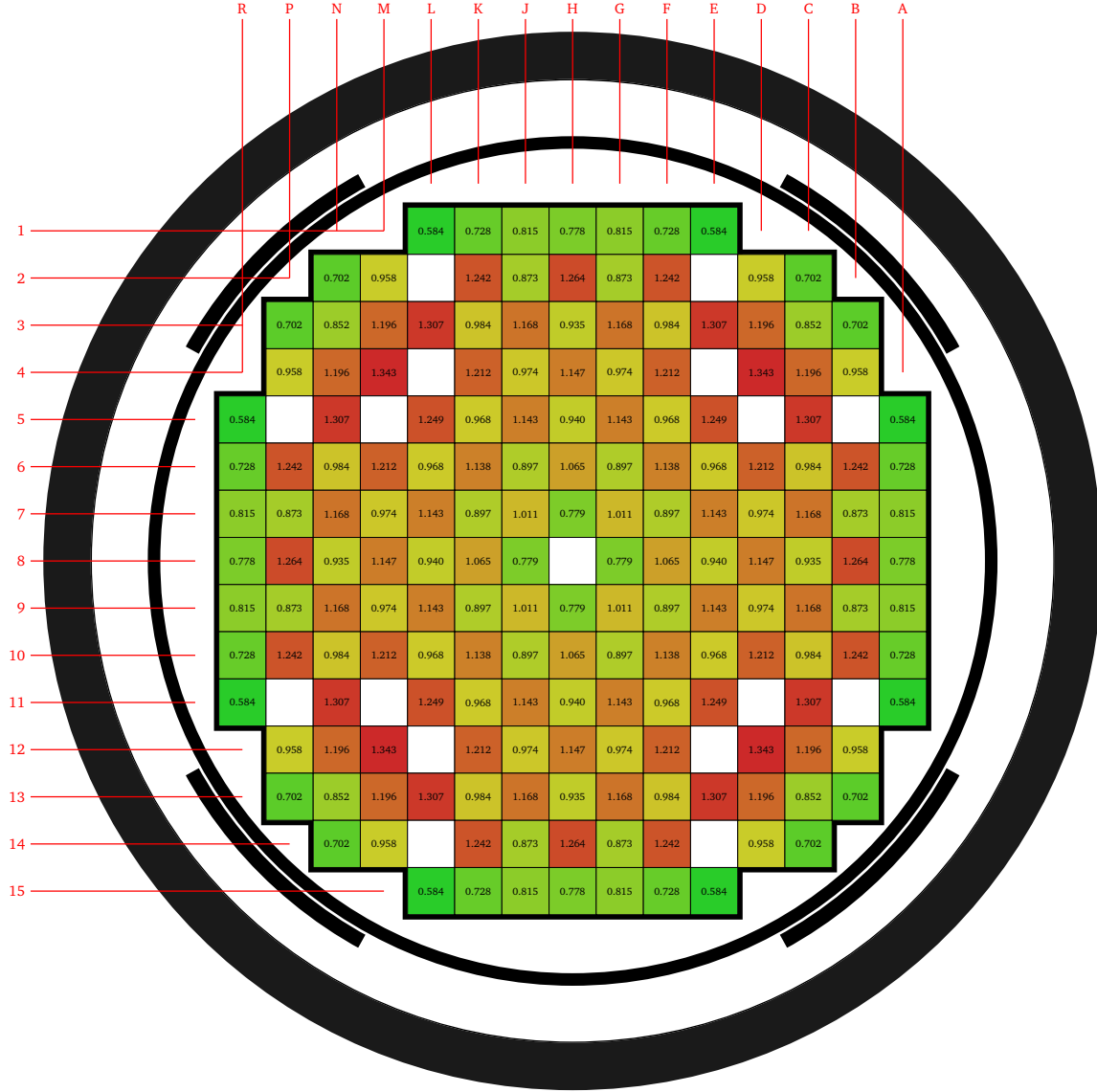


Figure 33: Radial detector measurements at HZP after linear tilt correction is applied. Note that all eighth-symmetric positions are extrapolated from an eighth-core symmetric radial map.

shifting of bundles with flow. This could explain the presence of a large tilt at HZP that dissipates with burnup as observed in the BEAVRS benchmark. To test this hypothesis, a water gap test was conducted by Sykora to determine whether adding a 0.5 cm water gap in the southeast corner between the baffle and assemblies would mirror the observed tilt in BEAVRS data [8]. The results from this test are included in Section 7.1.2.

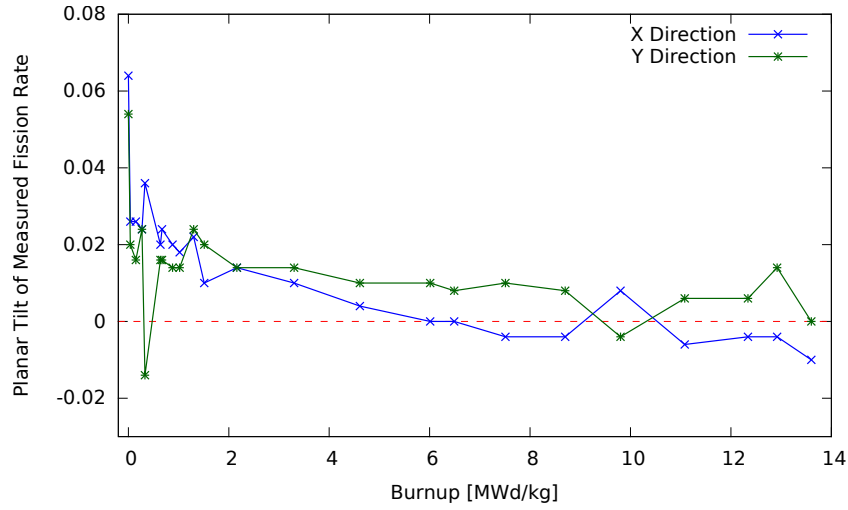


Figure 34: Magnitude of x and y tilts vs burnup for cycle 1.

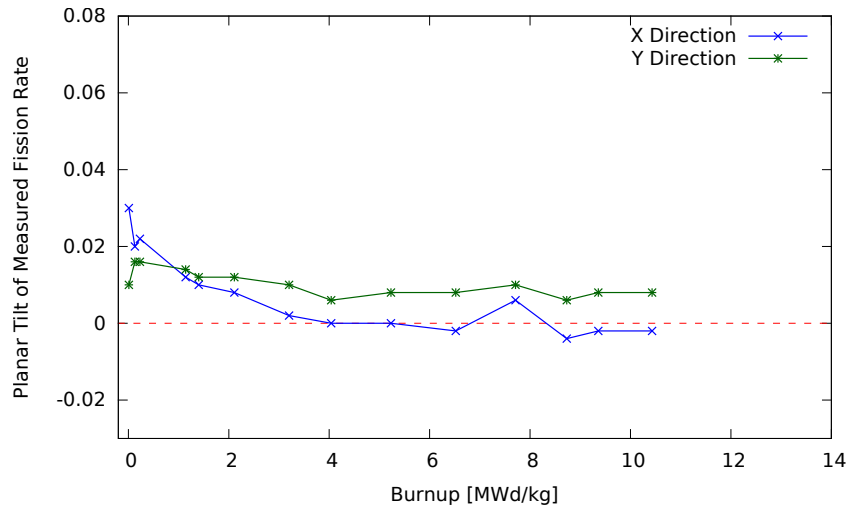


Figure 35: Magnitude of x and y tilts vs burnup for cycle 2.

7.1.2 Initial Gap Test to Mimic Tilt in BEAVRS Data

A 0.5 cm water gap can be modeled in SIMULATE between the fuel assemblies and the baffle in the southeast corner of the core. To illustrate the effect of an updated model which includes this water gap on overall modeling accuracy, a radial error map before and after the water gap inclusion at HZP is shown in Figures 36 and 37 respectively¹. While the core-wide RMS reduces

¹The radial error map shown in Figure 37 after adding a 0.5 cm water gap differs slightly from the one found in Sykora's Master's thesis because of updated BEAVRS specifications since the publishing of his findings. Further

dramatically, this simple water gap model does not wholly explain the cause of the tilt in BEAVRS data. A major reason for this is that the addition of the water gap induces an exponential tilt which helps reduce the errors on the periphery of the core but still adds significant errors in the core center, where a linear tilt is more suitable than an exponential one. Thus, accounting for varying water gaps throughout the core and not just at the edge will allow for tuning of the water gap to compensate for this phenomenon. Furthermore, Sykora finds that keeping a fixed water gap through depletion results in higher errors due to the resulting exponential tilt over-correcting the real tilt in BEAVRS data [8]. This suggests that the gap must vanish or reduce with burnup. A key element in tilt-correction uncertainty quantification therefore is to introduce more degrees of freedom in water gap distribution to reduce RMS and to match the BEAVRS radial map more closely to simulated results. This approach will be pursued in Section 7.1.3.

7.1.3 Methodology for Tilt-Correction Uncertainty Quantification

In order to improve Sykora’s water gap test, now water gaps between adjacent assemblies will also be varied to induce a tilt in SIMULATE that further replicates the one observed in BEAVRS data at HZP. Doing so hopefully will counteract the effect of the exponential tilt seen by varying just the reflector-assembly gaps. The presence of different water gaps throughout the core would affect moderation within each assembly and provide additional parameters to bring the tilt generated in SIMULATE closer to that detected in BEAVRS data. Thus, the task of tilt-correction uncertainty quantification is formulated into a large-scale optimization problem, whereby the objective function is a minimum RMS between BEAVRS data (not tilt corrected) and SIMULATE reaction rate data, and the optimization parameters within SIMULATE are the water gaps between adjacent assemblies as well as the water gap between outer assemblies and the baffle. Since the generation of SIMULATE reaction rate maps don’t follow a simple mathematical operator, traditional optimization methods that guarantee the location of a global minimum cannot be used. Furthermore, since there are a large number of inter-assembly gaps and assembly-baffle gaps that can be modeled, a brute-force approach of testing all possible gap combinations, even for a narrow range of gap perturbations, is highly unfeasible.

For this process, the parameters for 129 fuel regions and 64 reflector regions are perturbed². Water gaps between outer assemblies and the baffle can take on any values between 0.00 cm to 0.85 cm, in 0.01 cm increments, while the inter-assembly pitch is varied between 21.42 cm to 21.58 cm, in 0.01 cm increments. 21.42 cm represents the lower bound where there is no inter-assembly water gap, and 0.85 cm and 21.58 cm are chosen empirically based on the fact that any water gaps larger than these amounts have negligible effects on reducing RMS for any single assembly.

Initial testing focused on perturbing just the gaps in the reflector region to reduce RMS as much as possible without altering inter-assembly pitches. All reflector regions were grouped into octants and perturbed with very coarse parameter intervals to reduce computational constraints for an optimization problem that takes exponential-time to compute. It was found that lower RMS values

discrepancies also arise due to the fact that the SIMULATE-3 nodal diffusion solution was used instead of the CASMO-5 MxN full-core transport solution used by Sykora.

²All fuel assemblies with enrichments of 1.6% and 2.4% are perturbed. 3.1% enriched assemblies maintain a fixed assembly pitch of 21.50 cm. These assemblies are not included as part of the optimization simply to reduce the number of optimization parameters and also because it is believed that perturbations in the reflector region water gaps can compensate for any large errors in the 3.1% enriched assembly regions, since these assemblies are all located on the core periphery. Moreover, SIMULATE discretizes the reflector region into 64 nodes, adding a larger amount of reflector material where the neutron shield panel is present.

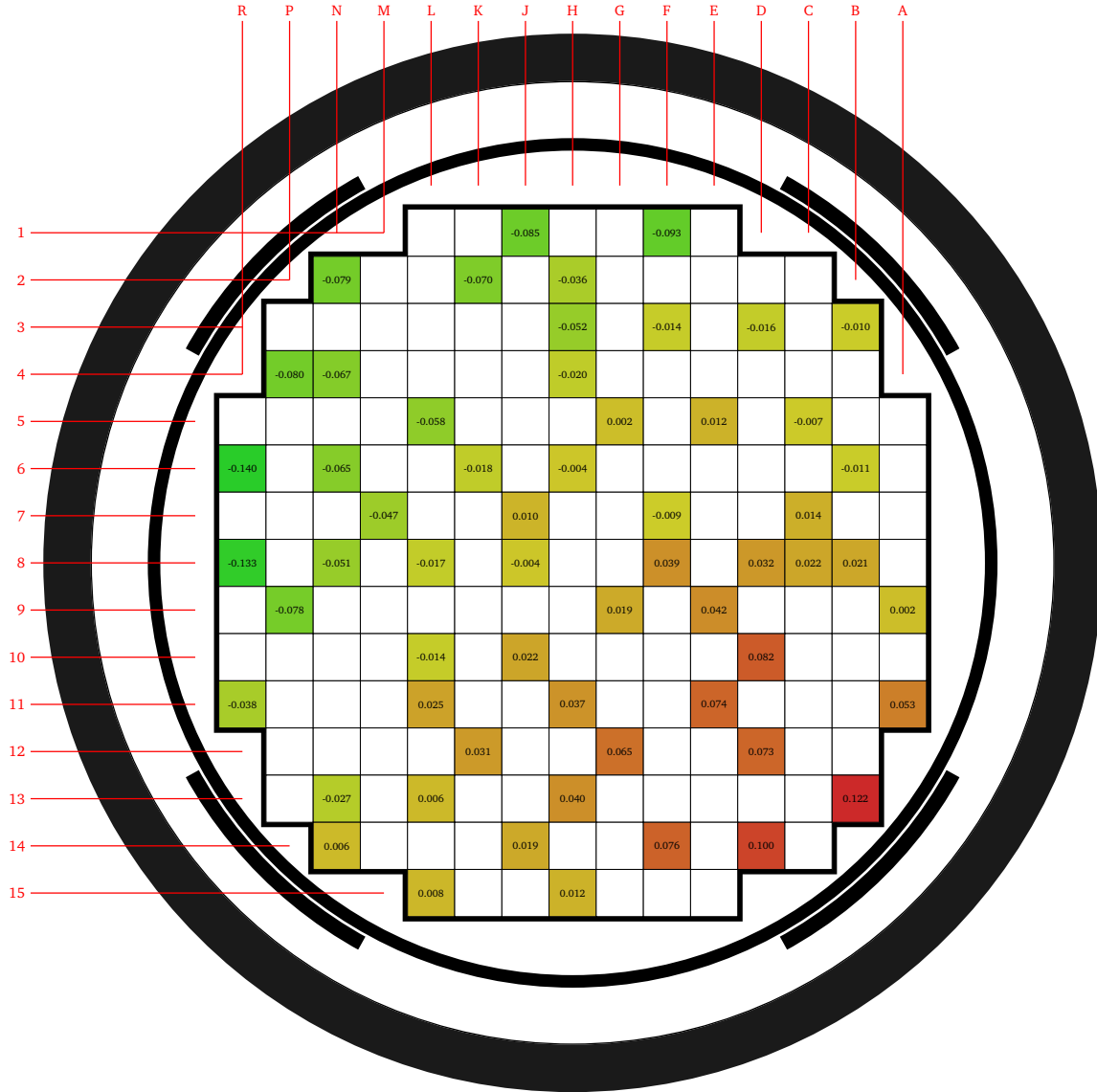


Figure 36: Radial error map between BEAVRS data without tilt-correction and CASMO/Simulate. RMS = 0.054.

were attained when gaps were 0.08 cm in the northwest region, 0.80 cm in the southwest region, and 0.22 cm in southwest and northeast regions. These results are consistent with Sykora's initial water gap test, which found that large water gaps in the southeast reflector regions and moderate water gaps in other quadrants tended to lower RMS.

Now, with this approximate water gap distribution in the reflector region, it is time to focus on the fuel regions. Given the sheer number of fuel regions that can be perturbed, it is now necessary

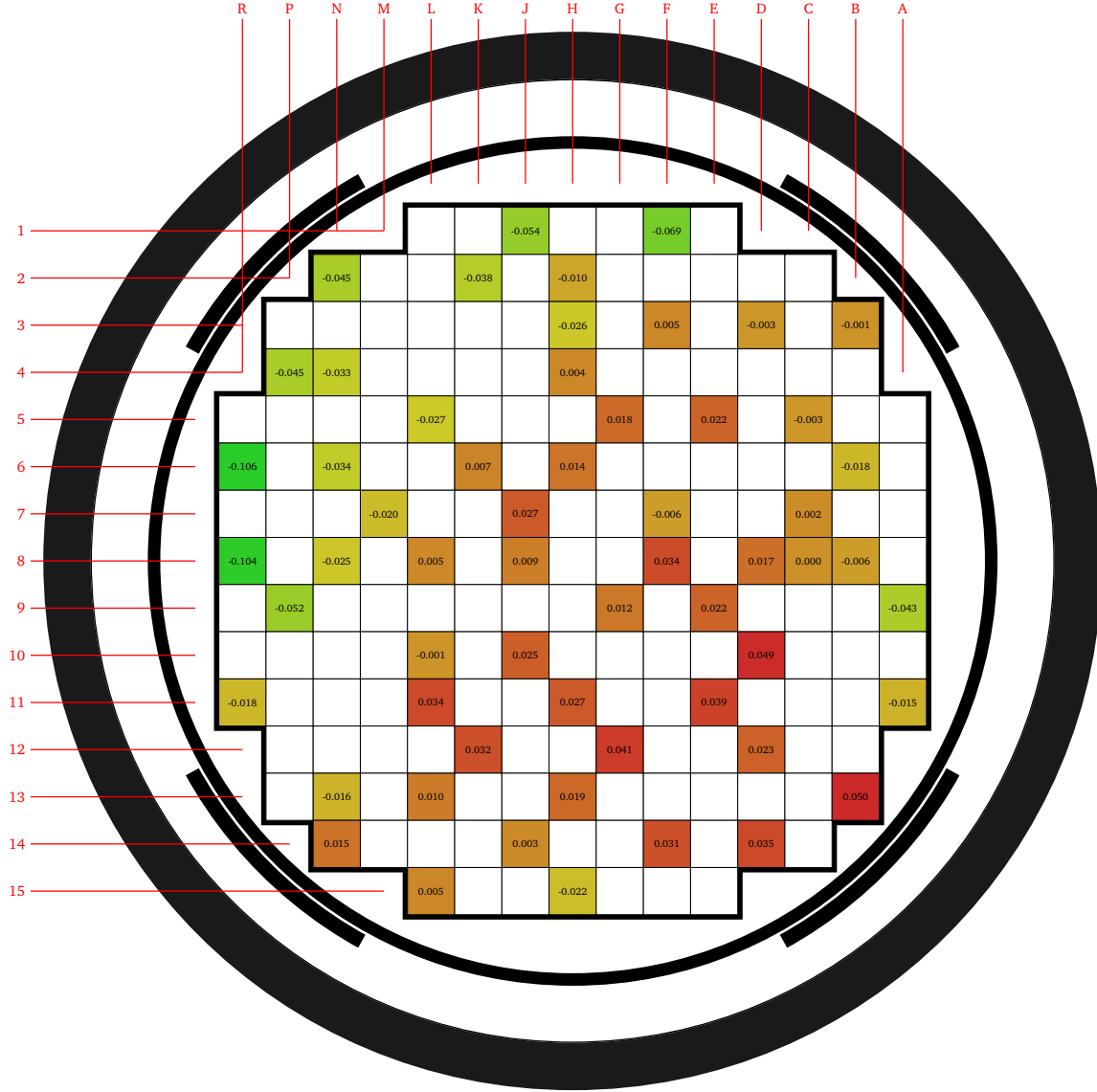


Figure 37: Radial error map between BEAVRS data without tilt-correction and CASMO/Simulate with the inclusion of a 0.5cm water gap in southeast corner. RMS = 0.034.

to impose additional constraints and changes to the optimization to reduce runtime from one that is exponential to one that is polynomial in nature. In order to achieve this, an initial gap distribution is selected. At each iteration, each of the 129 assemblies is perturbed by ten varying water gaps. The gap that minimizes RMS for that assembly is stored. After a single assembly perturbation, the gap is restored to its initial gap distribution before moving on to the next assembly to perturb. In this manner, the single assembly gap perturbation that leads to a minimum RMS over all 129 assemblies

is found and fixed for the remainder of the algorithm. In the subsequent iteration, 128 assemblies are now perturbed and the single assembly gap perturbation that minimizes RMS is found and fixed for that specific assembly. This process is repeated for all remaining unfixed assemblies until there are no more remaining assemblies to perturb or the RMS does not reduce from the previous iteration to the current iteration. Suppose each assembly can be initially perturbed by 10 different possible water gaps. Then a simple upper bound calculation for amount of time it would take for this entire optimization would be,

$$\begin{aligned}
\text{Total SIMULATE Runs} &= \left(129 \text{ Perturbed Assemblies} * \frac{10 \text{ Perturbations}}{\text{Assembly}} \right) + \\
&\quad \left(128 \text{ Perturbed Assemblies} * \frac{10 \text{ Perturbations}}{\text{Assembly}} \right) + \dots + \\
&\quad \left(2 \text{ Perturbed Assemblies} * \frac{10 \text{ Perturbations}}{\text{Assembly}} \right) + \\
&\quad \left(1 \text{ Perturbed Assembly} * \frac{10 \text{ Perturbations}}{\text{Assembly}} \right) \\
&= 83,850 \text{ SIMULATE Runs} \\
\text{Total Runtime} &= 83,850 \text{ SIMULATE Runs} * \frac{3 \text{ Seconds}}{\text{SIMULATE Run}} \\
&= 2.91 \text{ Days}
\end{aligned} \tag{14}$$

where this case would only be achieved if all gaps of the initial gap distribution were changed and RMS kept decreasing in each iteration. This is a much favorable alternative to testing all possible combinations of gap distributions within the 129 assemblies. In the latter case, this becomes an exponential problem with 10^{129} possibilities. A detailed description of the algorithm used to optimize the water gap distribution for a certain number of nodal regions is shown in Algorithm 1. It should be noted that the term **varying** refers to a Boolean datatype specifying whether a node should be perturbed or whether it will have a fixed gap for the remainder of the optimization.

Once an optimized water gap distribution results from perturbing the 129 fuel assemblies, the focus is shifted back to the reflector nodes. In the process of improving the water gap distribution within the assemblies, the reflector gaps were kept fixed, but now the reflector gaps can be varied while keeping the optimized fuel assembly gap distribution from the previous optimization fixed to attain a new minimum RMS. The same polynomial-time approach as Algorithm 1 is used to optimize the water gap distribution within the reflector nodes. Iterating back and forth between varying just the fuel assembly nodes and varying just the reflector nodes, a final optimized error map is generated, where now perturbing the gap in any single nodal region does not lower RMS any further. This error map is shown in Figure 38, which has an RMS of 0.7%. Beyond this point, the errors are very sensitive to a single assembly perturbation. Moreover, SIMULATE requires that there can only be 200 unique nodal segments, so this inhibits further optimization using gap perturbation increments finer than 0.01 cm.

It should be noted that while opting for a polynomial-time optimization over an exponential-time optimization saves tremendous runtime while exploring a wide range of perturbation combinations, a key shortcoming of this work is that this approach does not guarantee a global minimum. The implicit assumption for such an iterative scheme is that each successive perturbation could drive the

Algorithm 1 Algorithm for Optimizing Water Gap Distribution Within SIMULATE

```
1: Assign reference gap distribution ▷ Could come from a previous optimization
2:  $\text{RMS}_{\text{base}} \leftarrow \text{RMS}$  from reference gap distribution
3: Set all nodes that will be perturbed as varying
4: Assign all gap perturbations for each varying node
5: while At least one node is varying do
6:    $\text{RMS}_{\text{curr}} \leftarrow \text{RMS}_{\text{base}}$  ▷ Minimum RMS from current iteration
7:   for all node in nodes do
8:     if node is varying then
9:       Retrieve all gap perturbations for node
10:      for all gap perturbation in gap perturbations do
11:        Set node gap to gap perturbation
12:        Run SIMULATE with this node gap
13:         $\text{RMS}_{\text{temp}} \leftarrow \text{RMS}$  from this SIMULATE run
14:        if  $\text{RMS}_{\text{temp}} < \text{RMS}_{\text{curr}}$  then ▷ New minimum RMS found
15:          Store node and gap perturbation
16:           $\text{RMS}_{\text{curr}} \leftarrow \text{RMS}_{\text{temp}}$ 
17:        end if
18:      end for
19:      Reset node gap to value from reference case ▷ See Step 1
20:    end if
21:  end for
22:  if  $\text{RMS}_{\text{curr}} = \text{RMS}_{\text{base}}$  then ▷ No new minimum RMS found in iteration
23:    Break from while loop
24:  else ▷ Update minimum RMS and fix node gap
25:    Set node that yields  $\text{RMS}_{\text{curr}}$  as fixed ▷ No longer perturb this node
26:    Set node gap to perturbation value that yields  $\text{RMS}_{\text{curr}}$ 
27:     $\text{RMS}_{\text{base}} \leftarrow \text{RMS}_{\text{curr}}$ 
28:  end if
29: end while
```

final solution closer to a local minimum rather than a global minimum. However, for the purposes of this work, achieving a global minimum is not absolutely critical since the local minimum that was found in Figure 38 was within the bounds of a 1-sigma deviation when calculating the flux map uncertainties [5].

It should be noted once again here that a key assumption in the generation of this radial map is that predictive models are capable of capturing reactor behavior for short intervals. The closer that these models can be brought to mimic BEAVRS data, the stronger the argument that the behavior of operating data is accurately being modeled by simulation tools. Since the exact source of the tilt in BEAVRS data is not known and there is relative lack of certainty in the inter-assembly gap due to imperfect core loading, incorporating such gap variations for modeling purposes could be justifiable, especially given that the magnitude of these gap perturbations is quite minute. Furthermore, analysis is restricted to quantifying the uncertainty of the linear tilt at HZP, where the magnitude of this tilt is largest and represents a conservative upper-bound in uncertainty levels.

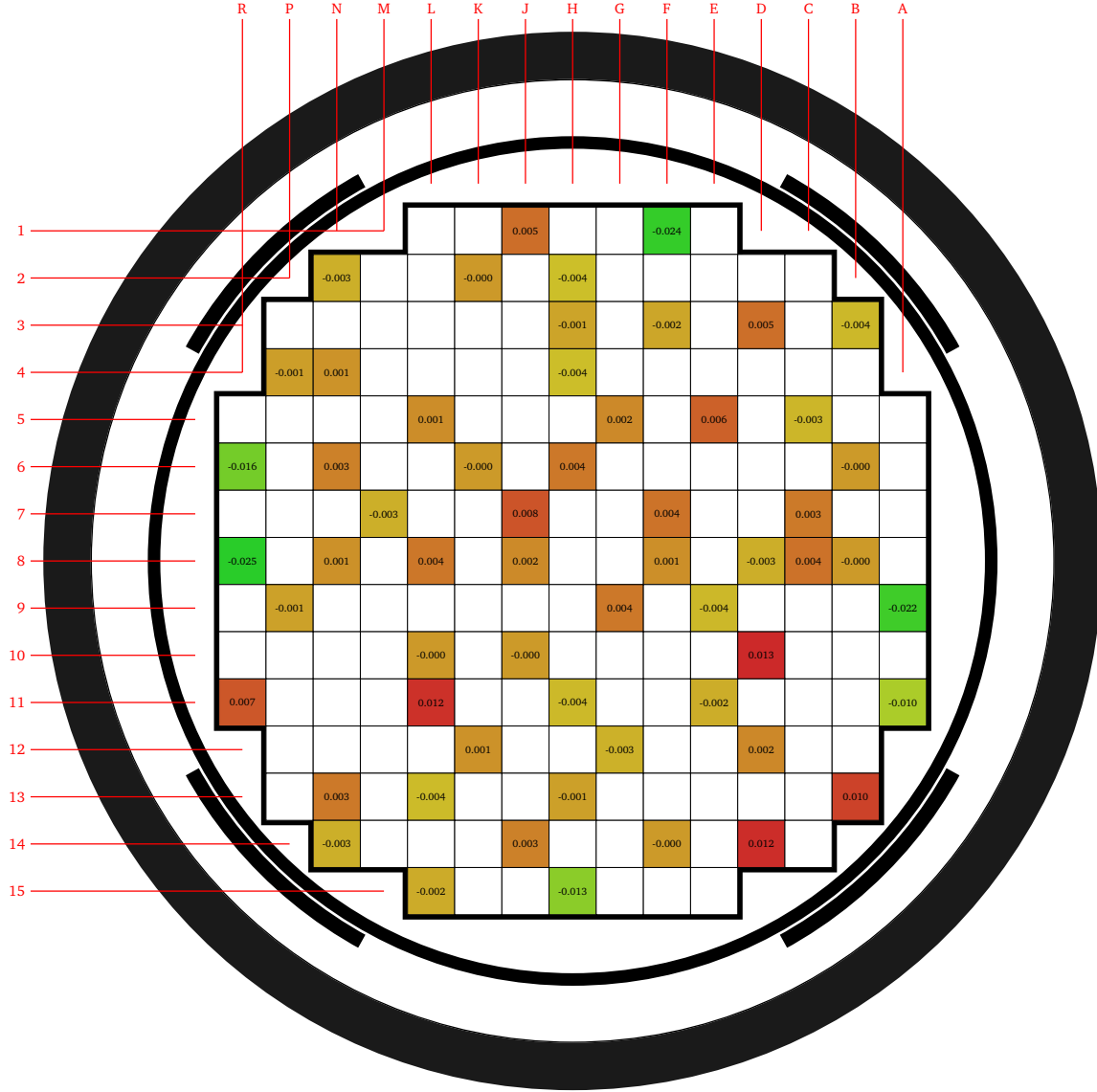


Figure 38: Radial error map between BEAVRS data without tilt-correction and CASMO/Simulate with optimized water gap distribution in baffle region and between adjacent assemblies. RMS = 0.007.

7.1.4 Evaluation of Tilt-Correction Uncertainty

With this tilt-induced Simulate map generated from a varying water gap distribution, there now exist two sets of comparable data between BEAVRS and Simulate to assess the added uncertainty of imposing a linear tilt correction to the data sets. The following procedure will be used to quantify

such uncertainty:

- (1) The linear planar tilt correction is applied to the BEAVRS case at HZP for cycle 1. The resulting map is shown in Figure 33.
- (2) Likewise, the linear planar tilt correction is applied to the optimized Simulate map generated by perturbing the water gap profile at HZP. This map represents the reference case that tilt-corrected HZP BEAVRS data will be compared to.
- (3) The resulting radial maps from Steps (1) and (2) are now octant core-symmetric. The two radial maps are collapsed to an eighth-core and relative errors at each octant-symmetric position are calculated to assess the underlying uncertainty in using tilt-corrected BEAVRS data for each assembly.

Following this procedure, the eighth-core collapsed error maps generated after Step (3) is shown in Figure 39. This map represents the resulting error in each assembly after imposing a linear tilt correction on both BEAVRS data as well as Simulate data with an induced gap that mimics the tilt in BEAVRS data.

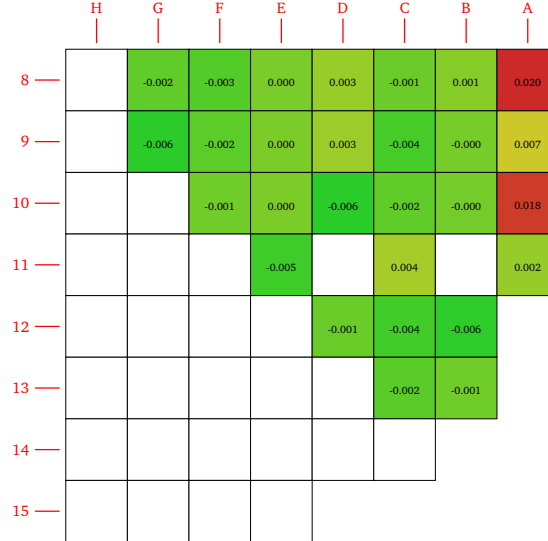


Figure 39: Radial error map after applying linear tilt correction to BEAVRS data and to CASMO/Simulate with the inclusion of a 0.5cm water gap in southeast corner. This map is eighth-core symmetric and represents a conservative upper-bound for the error in each assembly when using tilt-corrected data to model time-dependent BEAVRS data. RMS = 0.006.

Thus, when computing overall time-dependent uncertainty for each assembly by imposing a time-dependent model to BEAVRS data, Equation 15 will be used,

$$\left(\frac{\delta\phi}{\phi}\right)_{ij,time-dep} = \sqrt{\left(\frac{\delta\phi}{\phi}\right)_{ij,detection}^2 + \left(\frac{\delta\phi}{\phi}\right)_{ij,tilt-corr}^2} \quad (15)$$

where $(\delta_\phi/\phi)_{ij,time-dep}$ represents the time-dependent uncertainty of each radial assembly position, $(\delta_\phi/\phi)_{ij,detection}$ represents the error induced from fitting a simulation model to assembly-level BEAVRS data, and $(\delta_\phi/\phi)_{ij,tilt-corr}$ represents the conservative upper-bound error within each assembly that stems from assuming a linear tilt-correction to BEAVRS data. In other words, overall time-dependent uncertainty has a contribution from error that is rooted within assuming a model to fit time-dependent data as well as from an un-modeled phenomenon that accounts for an imperfect replication of the tilt observed within BEAVRS data. Section 7.2 focuses on the CASMO/Simulate model to fit BEAVRS data to and outlines a methodology for calculating $(\delta_\phi/\phi)_{ij,detection}$. The value of $(\delta_\phi/\phi)_{ij,tilt-corr}$ for each assembly, on the other hand, is given by the errors in Figure 39.

While the work in this section could be extended to multiple burnups by accounting for depletion, this effort will not be examined simply because there is no way of knowing how the water gaps vary with time. Each burnup step provides even more degrees of freedom to optimize over, making the problem even more intractable. There is no basis from operational data to confirm depletion behavior of these gaps, and thus strong assertions about the nature of the tilt in BEAVRS data with respect to time cannot be made. HZP is chosen as the focus for this section mainly because the tilt is highest at this point and thus the uncertainty introduced by imposing a tilt-correction would also result in the most conservative value.

7.2 Uncertainty in Using the CASMO/Simulate Model

The motivation for computing time-dependent uncertainty can be understood by seeing how operational data behaves over time. Assuming that at each burnup point, there is a constant but independent uncertainty section with each axially-integrated reaction rate for a particular assembly, the goal for this section is to figure out whether the time-varying behavior of reaction rates for each assembly follow an observable trend. If so, this trend is characterized and any deviation from this behavior can be attributed to an additional measure of time-dependent uncertainty. To this end, simulated reaction rates from CASMO-5 [9] lattice codes and Simulate-3 [10] nodal diffusion simulator will serve as an additional basis for comparison. Trends in two cycles of operation are considered, and as explained in Section 7.1, tilt-corrected BEAVRS data is used to fit linear trends.

Analysis from CASMO/Simulate is suitable, however, only when reactors are operating near steady-state conditions. Thus, the time-dependent analysis for this section is restricted to looking exclusively at burnup points that are full-power points, defined as those points within the cycle power history above 90% rated power. Figures 40 and 41 show the power history for cycle 1 and 2 respectively, with the green points representing burnups that are full power points. This entire burnup range of full power points will be subdivided into smaller subset of burnups, as the highly volatile power history makes it difficult to model reaction rates effectively over a very long burnup.

Section 7.2.1 begins by examining what core-wide RMS between Simulate and BEAVRS data looks like in order to provide a holistic perspective of how errors behave over burnup. This transitions into a narrower analysis of assembly-level behavior in Section 7.2.2, where it is found that simulations overpredict reaction rates by a small but persistent discrepancy labeled as model bias. The remaining detection uncertainty after correcting for model bias is evaluated over the entire burnup range in Section 7.2.3 in an attempt to produce a core-level error heuristic over the entire cycle. Section 7.2.4 tests the assumptions made during this analysis, while Section 7.2.5 verifies that the results are not dependent on the model used by running the same tests with data from the MPACT model rather than the CASMO/Simulate model. Data from this model was provided by the CASL team [11]. Finally, section 7.2.6 combines detection uncertainty with uncertainty from

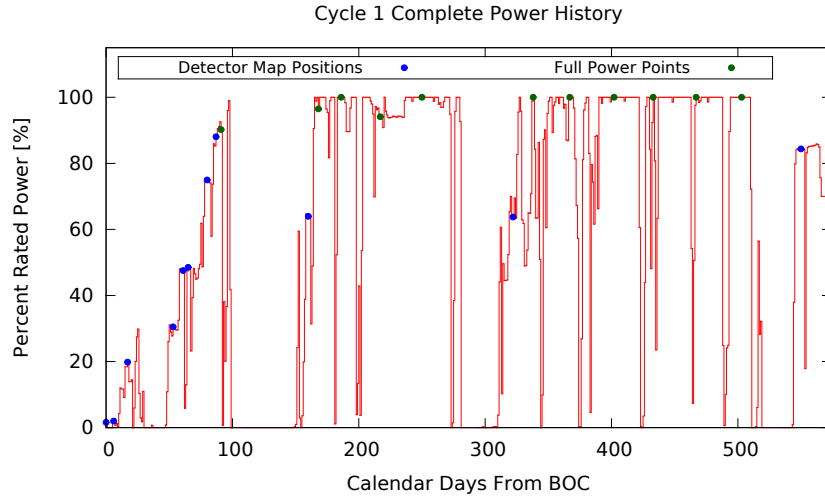


Figure 40: Cycle 1 burnup points. Full power points are indicated by green points (above 90% power). The list of full power points (in MWd/kg) are 1.02, 1.51, 2.16, 3.30, 4.61, 6.49, 7.51, 8.70, 9.80, 11.08, 12.34.

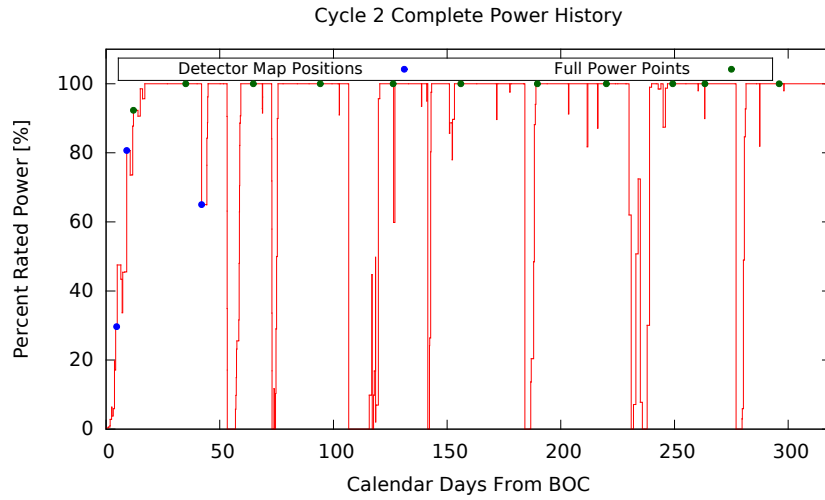


Figure 41: Cycle 2 burnup points. Full power points are indicated by green points (above 90% power). The list of full power points (in MWd/kg) are 0.23, 1.14, 2.11, 3.20, 4.04, 5.23, 6.52, 7.71, 8.73, 9.36, 10.43.

using tilt-corrected data to find overall time-dependent uncertainty.

7.2.1 Visualization of Core-wide RMS

Time series analysis can be conducted at the core level by observing how RMS between BEAVRS and Simulate maps vary at each burnup point. This is shown in Figure 42 for cycle 1 and in Figure 43 for cycle 2. From these graphs it is evident that core-wide RMS essentially decreases over burnup for both cases, and that the data for cycle 2 is much smoother than for cycle 1 due to the smoother power history in cycle 2. For cycle 1, there is reason to believe that burnup points 1.02 MWd/kg and 9.80 MWd/kg are outliers. The measurement readings at 9.80 MWd/kg happened during temporary shutdowns and sudden fluctuations, while the recordings at 1.02 MWd/kg occurred while the reactor was still ramping up and had not yet attained full power. These observations will become especially important in Section 7.2.4, where it is shown that analysis conducted with the outlier burnup points in cycle 1 yields erratic results since the state of the core is not well characterized during these transition periods.

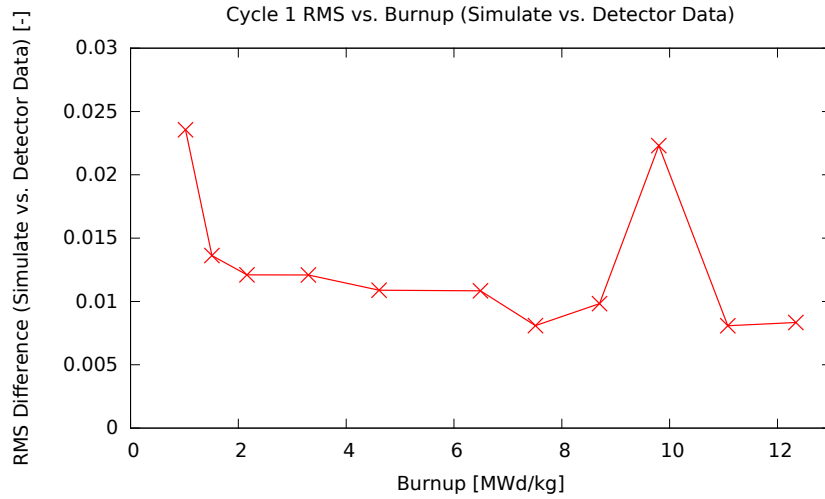


Figure 42: Core-wide normalized U-235 rate RMS error between CASMO-5/Simulate-3 and tilt-corrected BEAVRS data for cycle 1.

7.2.2 Characterization of Model Bias

The focus will now shift to assembly-level CASMO/Simulate data. As aforementioned, only a subset of burnups are going to be examined at a time, due to the unstable power history. Figure 44 plots reaction rates for CASMO/Simulate and BEAVRS data for assembly D10 at the first three burnups of 1.02, 1.51, and 2.16 MWd/kg in cycle 1. Looking at this region, it is clear that the two plots exhibit a similar shape but are offset by some amount. Defining this offset as model bias, the goal is to overlay the BEAVRS data onto the CASMO/Simulate data, and then measure how much noise there is between this new data set and CASMO/Simulate. This resulting noise can be quantified as a measure of detection uncertainty. Only assemblies that have a complete data set for these three burnups are used.

Thus, the methodology that is pursued is to calculate relative percent error at each burnup and

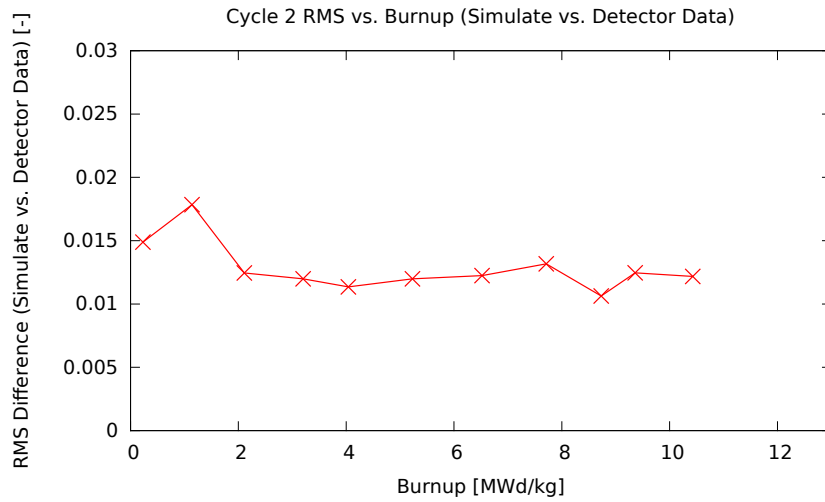


Figure 43: Core-wide normalized U-235 rate RMS error between CASMO-5/Simulate-3 and tilt-corrected BEAVRS data for cycle 2.

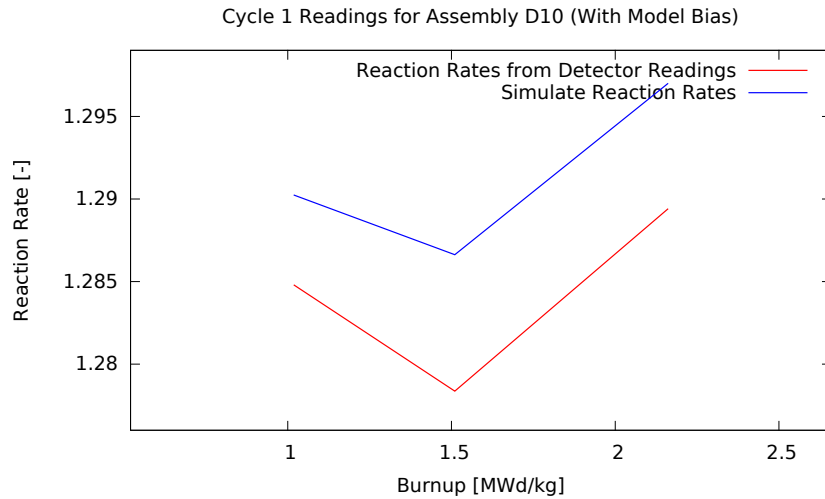


Figure 44: Reaction rates for CASMO/Simulate (blue) and BEAVRS (red) data for assembly D10 at 1.02, 1.51, and 2.16 MWd/kg.

average this value across all burnups to get an estimate of the average deviation between detector readings and simulated reaction rates. For each burnup, this average deviation is subtracted out and new detector readings are extrapolated based on this new percent error. Figure 45 shows the new extrapolated detector readings that are calculated once average relative error is subtracted out. The new BEAVRS data curve is basically a version of the original data that is overlaid onto

CASMO/Simulate data. With this new curve, RMS can be quantified and deemed as detection uncertainty. Model bias in this case is defined as the RMS value of percent errors between the original BEAVRS data and CASMO/Simulate data. For assembly D10, model bias is 0.6%, while detection uncertainty is 0.1%.

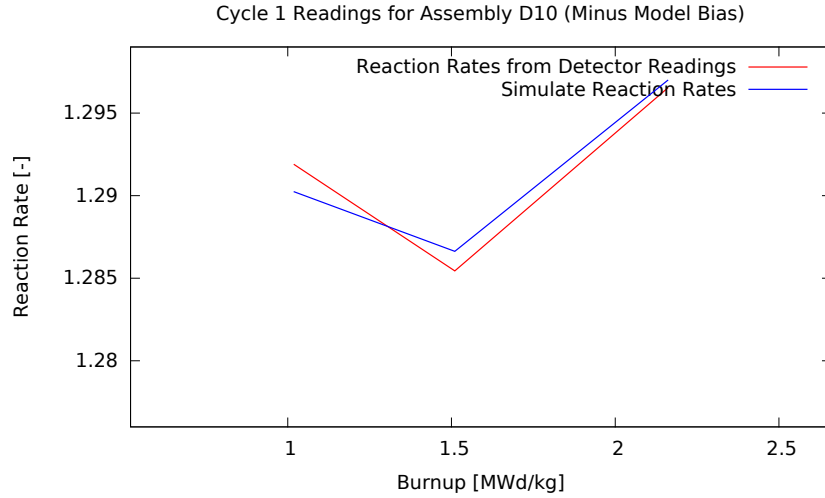


Figure 45: Reaction rates for CASMO/Simulate (blue) and BEAVRS data without model bias (red) for assembly D10 at 1.02, 1.51, and 2.16 MWd/kg.

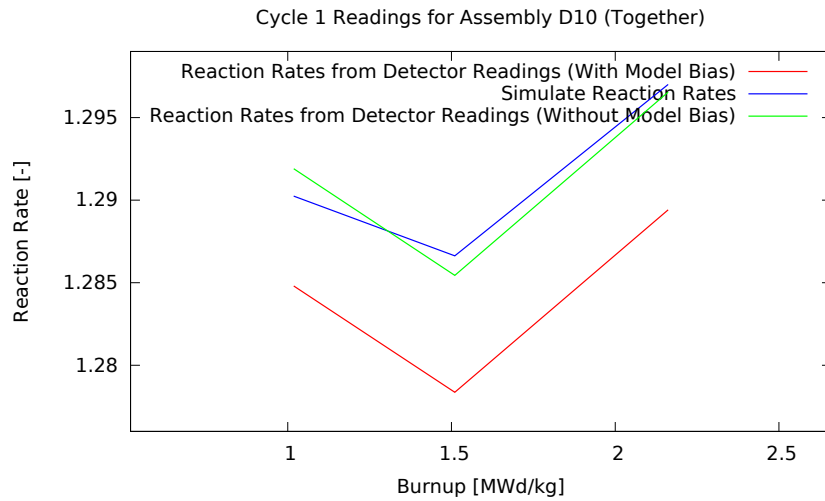


Figure 46: Reaction rates for CASMO/Simulate (blue) and BEAVRS (red) data, along with BEAVRS data with model bias subtracted out (green) for assembly D10 at 1.02, 1.51, and 2.16 MWd/kg.

By extending this methodology for all assemblies within the burnup range of 1.02, 1.51, and 2.16 MWd/kg, both model bias and detection uncertainty can be quantified for all assemblies. Figure 47 shows the results for all assemblies within the symmetric eighth-core. Once again, an RMS value that weighs both model bias and detection uncertainty by the number of symmetric positions for that specific assembly is calculated. RMS for model bias for the entire core is 1.8%, while the RMS for detection uncertainty is 0.9%.

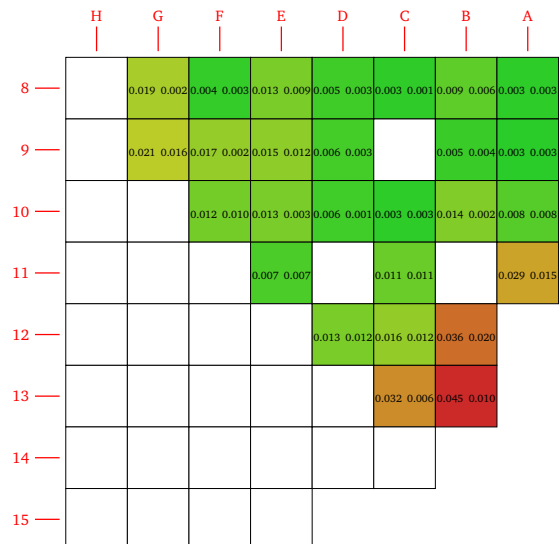


Figure 47: Cycle 1 model bias (top) and detection uncertainty (bottom) based on CASMO/Simulate model. Burnups: (1.02 1.51 2.16). Assembly-weighted model bias: 0.018. Assembly-weighted detection uncertainty: 0.009.

7.2.3 Quantifying Detection Uncertainty Over Entire Burnup Range

Assuming that burnup sets with non-overlapping points are independent of each other, Simulate data is now fitted for the remaining 8 burnup points that were not included in the first burnup set (1.02, 1.51, 2.16). The remaining 8 points are divided into two sets of four burnups so that Simulate data is compared to the burnup sets (3.3 4.61 6.49 7.51) and (8.7 9.8 11.08 12.34). The results for assembly-weighted core-wide model bias and detection uncertainty are summarized for all burnup ranges in Table 12. For cycle 2, the entire burnup range is divided up in a similar manner, where core model bias and detection uncertainties are computed for the burnup sets (0.23, 1.14, 2.11), (3.20, 4.04, 5.23, 6.52), and (7.71, 8.73, 9.36, 10.43). The results for cycle 2 are also given in Table 12.

Subdivision of the entire burnup range is carried out in such a way in order to extract detector variations from CASMO/Simulate data over a suitable time interval. Section 7.2.4 will look at the validity of burnup sets with non-overlapping points being independent of each other as well as how credible results are of being independent of burnup subdivision.

Now that BEAVRS data has been overlaid onto SIMULATE data to remove model bias, all remaining relative errors between BEAVRS data and SIMULATE data for each assembly over the

Cycle 1 Burnup Points	Core-Wide Model Bias (RMS)	Core-Wide Detection Uncertainty (RMS)
(1.02, 1.51, 2.16)	1.8%	0.9%
(3.30, 4.61, 6.49, 7.51)	1.1%	0.6%
(8.70, 9.80, 11.08, 12.34)	1.4%	0.9%
Cycle 2 Burnup Points	Core-Wide Model Bias (RMS)	Core-Wide Detection Uncertainty (RMS)
(0.23, 1.14, 2.11)	1.6%	0.6%
(3.20, 4.04, 5.23, 6.52)	1.2%	0.3%
(7.71, 8.73, 9.36, 10.43)	1.2%	0.5%

Table 12: Assembly-weighted model bias and detection uncertainty for all burnup sets in cycles 1 and 2.

three burnup divisions can be aggregated as a histogram to find a 95% confidence level uncertainty. Figure 48 plots errors between BEAVRS data and SIMULATE data for cycle 1, where the histogram frequencies are weighted by the number of symmetric assembly positions.

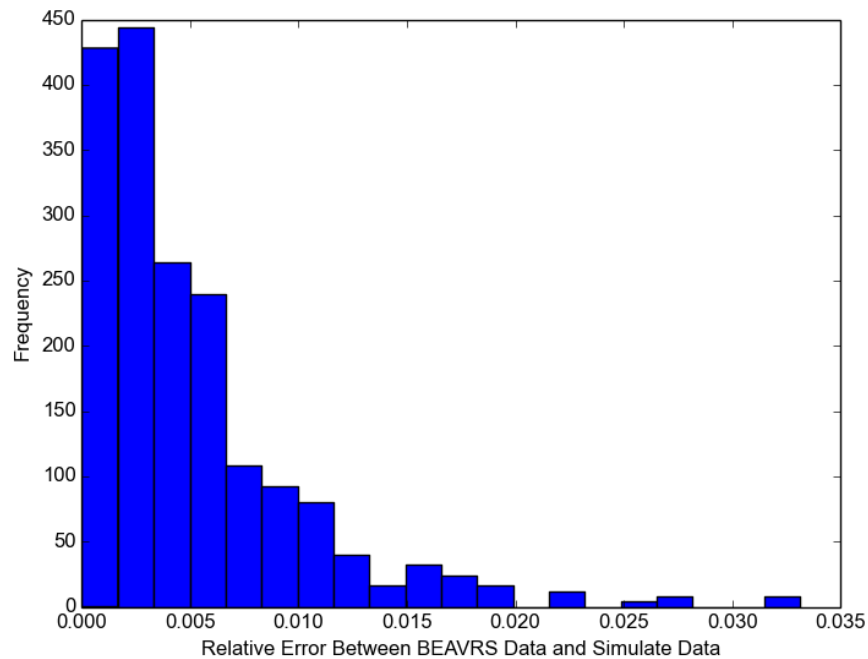


Figure 48: Histogram of relative error between Simulate Data and BEAVRS data after correcting for model bias within all assemblies over all burnups in cycle 1. RMS = 0.8%, 95% level = 1.6%.

From this histogram, it is clear that errors are skewed towards the lower end, with a few large

outliers on the higher end. The RMS value for detection uncertainty in cycle 1 is 0.8%, but taking a more conservative error estimate of 95% confidence yields a much higher detection uncertainty of 1.6% for cycle 1. Following a similar procedure for cycle 2 yields a 95% confidence value of 0.9% for detection uncertainty.

7.2.4 Validation of Assumptions for the CASMO/Simulate Model

There are two main assumptions that were made when calculating the results by using the CASMO/Simulate model. Firstly, the entire burnup range for both cycles were subdivided into three non-overlapping burnup ranges containing three or four burnups, and it needs to be shown that the results are not dependent on how the subdivisions were made. More specifically, for cycle 1, the entire burnup set (in MWd/kg) is subdivided into the three burnup ranges of (1.02, 1.51, 2.16), (3.30, 4.04, 5.23, 6.52), and (7.61, 8.73, 9.36, 10.43). However, the subdivision could have also been (1.02, 1.51, 2.16, 3.30), (4.04, 5.23, 6.52), and (7.61, 8.73, 9.36, 10.43). Thus, it needs to be shown that detection uncertainty results are independent of the burnup subdivision choice. Table 13 shows the results for the 95% confidence interval (CI) of detection uncertainty for all possible burnup subdivisions for both cycle 1 and cycle 2, and the results show that there is very little change in the results.

Cycle 1 Burnup Subdivision	95% CI of Detection Uncertainty
(1.02 1.51 2.16) - (3.30 4.61 6.49 7.51) - (8.70 9.80 11.08 12.34)	1.6%
(1.02 1.51 2.16 3.30) - (4.61 6.49 7.51) - (8.70 9.80 11.08 12.34)	1.7%
(1.02 1.51 2.16 3.30) - (4.61 6.49 7.51 8.70) - (9.80 11.08 12.34)	1.6%
Cycle 2 Burnup Subdivision	95% CI of Detection Uncertainty
(0.23 1.14 2.11) - (3.20 4.04 5.23 6.52) - (7.71 8.73 9.36 10.43)	0.9%
(0.23 1.14 2.11 3.20) - (4.04 5.23 6.52) - (7.71 8.73 9.36 10.43)	1.0%
(0.23 1.14 2.11 3.20) - (4.04 5.23 6.52 7.71) - (8.73 9.36 10.43)	0.9%

Table 13: Results of detection uncertainty based on three different ways of subdividing all burnups for each cycle.

The second assumption that needs to be validated that requires a bit more work is to show that non-overlapping burnup subdivisions within the entire burnup set are actually independent of each other. For example, if cycle 1 burnups are divided into the three ranges, (1.02, 1.51, 2.16), (3.30, 4.04, 5.23, 6.52), and (7.61, 8.73, 9.36, 10.43) as before, it needs to be shown that the three burnup ranges are independent of each other. To do this, the results of assembly-weighted detection uncertainty for all adjacent burnup choices of length 3 or 4 are looked at after accounting for model bias. If all non-overlapping burnup ranges of length 3 or 4 are truly independent of each other, then the results for detection uncertainty should remain the same regardless of how these burnups are chosen, so long as those burnups are adjacent to each other and represent a short interval for which predictive models are considered accurate. Table 14 summarizes the results for detection uncertainty for all possible adjacent burnup choices of length 3 in cycle 1 and 2, while Table 15 summarizes the results for burnup choices of length 4 in cycle 1 and 2.

From these two tables, it can be seen that RMS values of detection uncertainty and model

Cycle 1 Burnup Choice	Detection Uncertainty (RMS)	Cycle 2 Burnup Choice	Detection Uncertainty (RMS)
1.02 1.51 2.16	0.9%	0.23 1.14 2.11	0.6%
1.51 2.16 3.30	0.3%	1.14 2.11 3.2	0.5%
2.16 3.30 4.61	0.3%	2.11 3.2 4.04	0.4%
3.30 4.61 6.49	0.5%	3.2 4.04 5.23	0.3%
4.61 6.49 7.51	0.5%	4.04 5.23 6.52	0.3%
6.49 7.51 8.70	0.4%	5.23 6.52 7.71	0.5%
7.51 8.70 9.80	1.0%	6.52 7.71 8.73	0.5%
8.70 9.80 11.08	0.9%	7.71 8.73 9.36	0.5%
9.80 11.08 12.34	0.8%	8.73 9.36 10.43	0.3%

Table 14: RMS of detection uncertainty when looking at all combinations of burnup subsets of length 3 in cycle 1 and cycle 2. Outliers are marked in red.

Cycle 1 Burnup Choice	Detection Uncertainty (RMS)	Cycle 2 Burnup Choice	Detection Uncertainty (RMS)
1.02 1.51 2.16 3.30	0.8%	0.23 1.14 2.11 3.20	0.6%
1.51 2.16 3.30 4.61	0.4%	1.14 2.11 3.20 4.04	0.6%
2.16 3.30 4.61 6.49	0.5%	2.11 3.20 4.04 5.23	0.4%
3.30 4.61 6.49 7.51	0.5%	3.20 4.04 5.23 6.52	0.3%
4.61 6.49 7.51 8.70	0.5%	4.04 5.23 6.52 7.71	0.5%
6.49 7.51 8.70 9.80	1.0%	5.23 6.52 7.71 8.73	0.5%
7.51 8.70 9.80 11.08	1.0%	6.52 7.71 8.73 9.36	0.5%
8.70 9.80 11.08 12.34	0.9%	7.71 8.73 9.36 10.43	0.5%

Table 15: RMS of detection uncertainty when looking at all combinations of burnup subsets of length 4 in cycle 1 and cycle 2. Outliers are marked in red.

bias are roughly consistent for all burnups of length 3 and 4 in cycle 2, where RMS of detection uncertainty ranges from 0.3% to 0.6%. However, within cycle 1, there are some burnup subsets that have inflated RMS values totaling as much as 1.0%, and these entries are marked in red. In addition, the burnup points common to these high error points are also highlighted in red, and outliers are driving up uncertainty values. With the presence of these outlier burnup points, the errors are much higher than in other cases. With the presence of these outliers, the assumption that the results of non-overlapping burnup sets are independent does not hold true.

Recall that these highlighted burnups are the same ones that were found in Section 7.2.1 to have added model inaccuracies simply due to unstable operating conditions, further corroborating the existence of outliers at these two points. The simplest way to deal with these outliers is to remove them from the complete burnup range, redo the analysis, and see how the results change. Without these outliers, the detection uncertainties are lower and are also now consistent among all burnup choices for both cycles. The total remaining number of burnup points in the complete range after removing outliers in cycle 1 is now 9, so the burnup range is subdivided into 3 burnup sets of length three. There is no need for another sensitivity analysis based on how to subdivide

the entire burnup range as there is now only one way of subdividing the entire burnup range into subsets of length 3.

Cycle 1 95% CI of Detection Uncertainty (with outliers)	Cycle 1 95% CI of Detection Uncertainty (without outliers)
1.6%	0.8%

Table 16: Cycle 1 95% CI of detection uncertainty with and without outliers at 1.02 and 9.80 MWd/kg.

Cycle 1 Burnup Choice	Detection Uncertainty (RMS)
1.51 2.16 3.30	0.3%
2.16 3.30 4.61	0.3%
3.30 4.61 6.49	0.5%
4.61 6.49 7.51	0.5%
6.49 7.51 8.70	0.4%
7.51 8.70 11.08	0.4%
8.70 11.08 12.34	0.4%

Table 17: RMS of detection uncertainty when looking at all combinations of burnup subsets of length 3 in cycle 1 when outliers at 1.02 and 9.80 MWd/kg are removed.

Table 16 shows that the 95% CI of detection uncertainty in cycle 1 reduces from 1.6% to 0.8% by running the same CASMO/Simulate model without the outlier burnups. Similarly, without the outlier burnups there is now a more consistent level of detection uncertainty across all cycle 1 burnup choices of length three, as shown in Table 17. Thus, there is now reason to believe that without these outliers, each burnup set of length 3 or 4 is independent of each other, as long as the burnups under consideration are adjacent to each other. Cycle 2 does not contain any outlier burnup points, with agreeing results among all burnup subdivisions.

7.2.5 Verification of Results with Other Simulation Tools

To show that these results are also not dependent on the model being used, data from MPACT, an alternative neutronics code, is also used as a basis for comparison against BEAVRS data. MPACT is a 3-D full-core transport solver that utilizes the 2-D/1-D method to couple 2-D Method of Characteristics (MOC) solutions radially with lower order 1-D diffusion or P_3 solutions axially [12]. While MPACT provides axial data, only axially-integrated results will be used since CASMO/Simulate does not really provide axial distributions.

For the purposes of comparison, the same methodology used in Sections 7.2.2 through 7.2.4 to generate results analogous to those found in Tables 16 and 17, using MPACT data instead of data from CASMO/Simulate. The resulting calculations are shown in Tables 18 and 19, and there is

tremendously close agreement between the final errors after correcting for model bias. Note that only cycle 1 data was available with MPACT.

Cycle 1 95% Confidence Interval of Errors		
Model	Detection Uncertainty (with outliers)	Detection Uncertainty (without outliers)
CASMO/Simulate	1.6%	0.8%
MPACT	1.6%	0.8%

Table 18: Cycle 1 95% CI of detection uncertainty with and without outliers at 1.02 and 9.80 MWd/kg using both CASMO/Simulate and MPACT.

Cycle 1 Burnup Choice	Detection Uncertainty in Simulate (RMS)	Detection Uncertainty in MPACT (RMS)
1.51 2.16 3.30	0.3%	0.4%
2.16 3.30 4.61	0.3%	0.5%
3.30 4.61 6.49	0.5%	0.3%
4.61 6.49 7.51	0.5%	0.3%
6.49 7.51 8.70	0.4%	0.4%
7.51 8.70 11.08	0.4%	0.4%
8.70 11.08 12.34	0.4%	0.5%

Table 19: RMS of detection uncertainty when looking at all combinations of burnup subsets of length 3 in cycle 1 when outliers at 1.02 and 9.80 MWd/kg are removed in CASMO/Simulate and MPACT.

7.2.6 Combining Detection Uncertainty and Uncertainty from Using Tilt-Corrected Data

Throughout this report, the term detection uncertainty was used to indicate the amount of error that remained from using the CASMO/Simulate model after correcting for model bias. Plotting the different values of detection uncertainty across all assemblies and burnup subsets as a histogram similar to Figure 48 demonstrates a 95% confidence level for $(\delta_\phi/\phi)_{ij,detection}$ in Equation 15. However, this error level does not account for the additional error present due to the fact that tilt-corrected BEAVRS data was used. Equation 15 can be used to compute overall time-dependent uncertainty $(\delta_\phi/\phi)_{ij,time-dep}$, and plotting this updated time-dependent uncertainty value for all assemblies in all burnup subdivision as a single histogram yields a 95% confidence value of 1.8% for cycle 1 and 1.9% for cycle 2.

This suggests that if model bias could be completely corrected for, then the resulting time-dependent uncertainty due to fitting data from reactor analysis software to operational data is on the same order as axially integrated measurement uncertainty. The uncertainty that arises from measurement carries over throughout the core burnup and can be used to explain the random fluctuations that occur when trying to fit time-dependent trends to tilt-corrected BEAVRS data.

7.3 Uncertainty in Using a Linear Regression Model

In addition to using the CASMO/Simulate model to compute $(\delta_\phi/\phi)_{ij,detection}$ in Equation 15, this section looks at a linear model fit - an alternative but simpler time-dependent model - to compute the detection uncertainty. The reason for believing that simple linear fits would work well to explain time-dependent behavior of reaction rates hinges on the principle of the linear reactivity model, which states that at steady-state operations reactivity and hence reaction rates follow a linear trend with burnup [13]. This model is only valid for intervals where the reactor is running at ideal operating conditions with minor disturbances to core power. Thus, the time-dependent analysis for this section is once again restricted to looking exclusively at burnup points that are full-power points. Similar to section 7.2, the analysis will be restricted to using tilt-corrected BEAVRS data. This is because the tilts vary with time and produce additional biases that drive time-dependent data further from behaving linearly. Tests will also be conducted without the outlier points that were detected in Section 7.2.4 in order to generate comparable results as section 7.2. Section 7.3.1 presents a mathematical formulation and the requirements for fitting a linear regression to BEAVRS data. Section 7.3.2 tests these requirements, while Section 7.3.3 concludes by discussing the limitations of a linear model and calls for more sophisticated models that captures higher-order effects.

7.3.1 Formulation of the Linear Model

Models for time series analysis fall under two broad categories - stationary and non-stationary models. Stationary models assume that statistical properties such as mean, variance, and autocorrelation of the underlying distribution do not vary over time, while non-stationary models have statistical properties that are time-dependent.

Written mathematically, a stationary model can be expressed as:

$$y_t = \mu + \epsilon_t \quad (16)$$

and a non-stationary model can be expressed as:

$$y_t = \mu_t + \epsilon_t \quad (17)$$

where y_t is the regressed variable, ϵ_t can be viewed as the residual between the regressed and observed value, and μ is time-independent, while μ_t is time-dependent [14]. For a non-stationary linear regression, μ_t can be expressed as:

$$\mu_t = \beta_0 + \beta_1 t \quad (18)$$

For a linear regression model to be valid, however, it needs to meet four requirements [15]:

- (1) The linear regression exhibits a high value of R^2
- (2) Residuals of regressed values and measured values are independent of each other with respect to time.
- (3) Residuals of regressed values and measured values are homoscedastic.
- (4) Residuals of regressed values and measured values exhibit a normal distribution centered around a 0 mean.

All four of these conditions will be tested in the following section using a combination of visual and statistical techniques.

7.3.2 Verification of Requirements for a Linear Model

Homoscedasticity refers to the fact that all random variables in the sequence have the same finite variance, while R^2 is a measure of goodness of fit of a linear model to the data, and is a value between 0 and 1 that describes the percentage of variable variance that is explained by the linear model. A higher value of R^2 implies that a the model is indeed a good fit for the data, and is defined in this case mathematically as [15]:

$$R^2 = 1 - \frac{\sum_T \epsilon_t^2}{\sum_T (y_t - \bar{y})^2} \quad (19)$$

where \bar{y} is the mean observed value.

Requirement 1: High R^2 values for regressed lines

For each assembly, a linear trendline is fit to the time-dependent tilt-corrected BEAVRS data. When this is done, the R^2 value for this fitted model as well as the RMS can be calculated for each assembly. For example, Figure 49 shows BEAVRS data compared to a linear fit for assembly G9. For this specific assembly, R^2 is equal to 95.5% and the RMS is 0.9%. Figures 50 and 51 summarize both of these values as radial maps for cycle 1 data and cycle 2 data respectively, where the top number in each assembly corresponds to RMS error, while the bottom number corresponds to the R^2 goodness of fit metric. Since the core is eighth-symmetric and numerical results are exactly the same for all symmetric positions, only the results for an eighth-core are shown. These maps illustrate that assemblies with high values of R^2 typically have low RMS values.

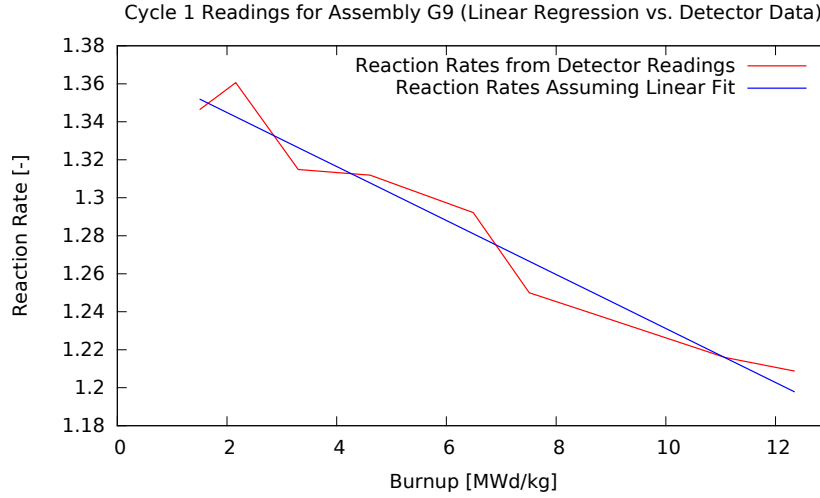


Figure 49: Cycle 1 measured and linearly regressed reaction rates vs. burnup for assembly G9.

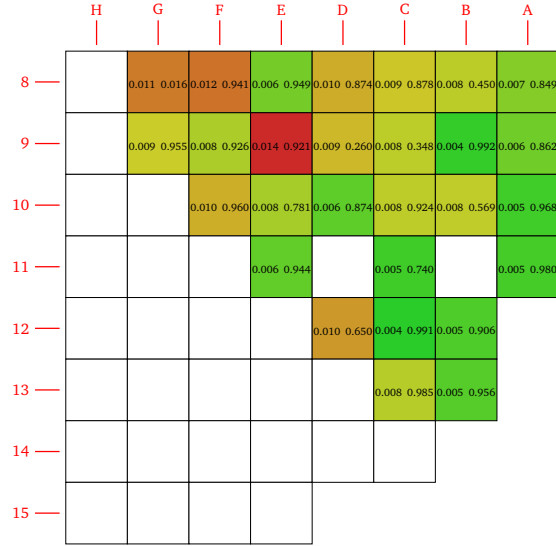


Figure 50: Cycle 1 RMS (top) and R-Squared (bottom) values based on applying linear model to data set of reaction rate vs. burnup for each assembly. Assembly-weighted total RMS = 0.008.

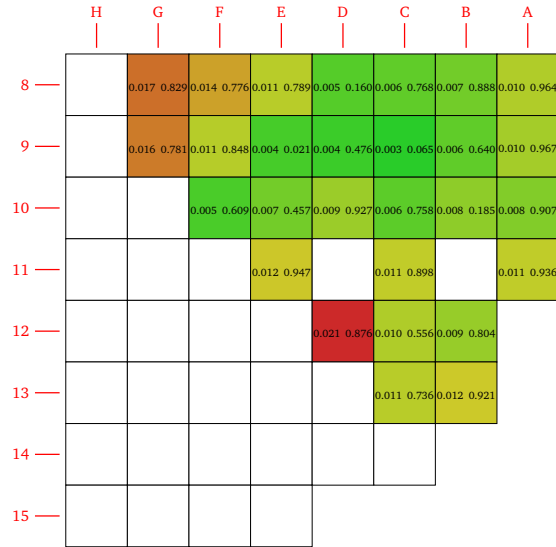


Figure 51: Cycle 2 RMS (top) and R-Squared (bottom) values based on applying linear model to data set of reaction rate vs. burnup for each assembly. Assembly-weighted total RMS = 0.010.

Requirement 2: Residuals are homoscedastic

As aforementioned, homoscedasticity of residuals refers to the fact that the variance of the residuals approach the same constant value. This means that the variance around the regression line must be roughly the same, and a cone-shaped pattern of data points distributed around the trendline would indicate heteroscedasticity [16]. Visually, a spread level plot of absolute studentized residuals versus regressed values should show the residuals be approximately equally distributed across the range of regressed values without any noticeable patterns [17]. Here, studentized residuals refer to the fact that the residuals have been divided by their standard deviation estimates for normalization purposes. Figure 52 shows the spread level plot for assembly G9, along with a linear trendline for the data to show how residuals behave as a function of regressed value. A flat trendline demonstrates homoscedasticity.

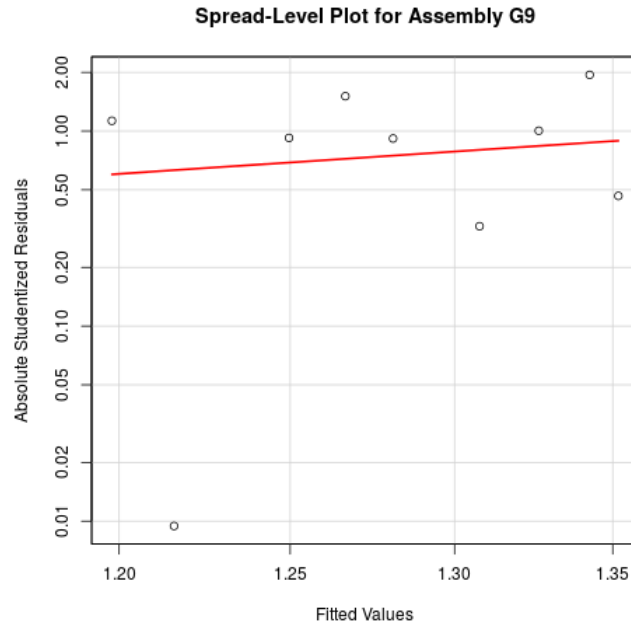


Figure 52: Spread level plot for linear model applied to for assembly G9 in cycle 1.

While visual tests are subject to individual interpretation as to whether residuals are homoscedastic, statistical tests can be used to test for the same criteria. The Breusch-Pagan test, developed by Trevor Breusch and Adrian Pagan and independently suggested by Dennis R. Cook and Sanford Weisberg, tests for the heteroscedasticity in a linear regression model using a test statistic based on the chi-squared distribution [18, 19]. The null hypothesis states that the residuals in the linear regression model under consideration are homoscedastic, and if the calculated p-value $p < \alpha$, where α is the statistical significance level, then there is evidence to reject the null hypothesis. For this thesis, α is chosen at a level of 0.05, so a p-value less than this value means that there is a 95% confidence level that the null hypothesis should be rejected. For the case of assembly G9, the calculated p-value is 0.56 so there is little evidence to suggest that the errors are

heteroscedastic. A complete list of p-values for the Breusch-Pagan test for each assembly regression is shown in Table 20. From this table, it can be seen that 25 out of 28 assemblies pass the statistical test, which suggests that errors are homoscedastic for most assembly regressions.

Requirement 3: Independent residuals with respect to time

Tests for independence of residuals are difficult to visualize, so this section will focus solely on statistical tests to determine whether residuals are correlated to each other. A common avenue for such analysis is to use the Durbin-Watson test for autocorrelation of disturbances [20, 17]. The null hypothesis states that the autocorrelation of the residuals is 0, implying independence of residuals. The test can be conducted for either positive correlation or negative correlation, so a two-sided test is conducted to test against the alternative hypothesis that the autocorrelation is not equal to 0. Once again, the null hypothesis is rejected for p-values less than 0.05, and Table 20 indicates that most assemblies pass this test at the 95% level, strongly suggesting that regressed residuals are independent from each other.

Requirement 4: Residuals exhibit a normal distribution centered around 0

Residuals being normally distributed can be shown both graphically and statistically. Visually, a normal Quantile-Quantile (Q-Q) plot can be generated for the residuals to show normality [17]. A Q-Q plot plots studentized residuals against theoretical quantiles of a normal distribution. The plotted points should fall along the line $y = x$, indicating that a specific quantile of residual lies exactly where the expected quantile from a normal distribution is. Figure 53 shows an example of this for assembly G9, with the red line indicating a best fit line for the plotted data.

Once again, statistical tests can be utilized to test for normality in a more definitive manner. The Shapiro-Wilk test tests the null hypothesis that a sample distribution comes from a normally distributed population [21, 22, 17]. It should be noted that the Shapiro-Wilk test does not assume the underlying sample distribution comes from a standard normal distribution, but instead tests for a general normal distribution with unknown population parameters. A complete list of p-values for each assembly when a linear regression is applied is shown once again in Table 20. For cycle 1, 27 out of the 28 assemblies resemble a normal distribution.

7.3.3 Limitations of the Linear Model

In general, analysis of the conditions for a linear regression model indicates that linear models are quite adequate for first-order fitting purposes. Assuming that a linear model is indeed valid and RMS values shown in Figures 50 and 51 are indicative of the detection error, $(\delta_\phi/\phi)_{ij,detection}$ found in Equation 15. A similar approach outlined in Section 7.2.6 can be followed that uses Equation 15 to find the 95% value for detection uncertainty. Using the linear model, it is found that the 95% CI for detection uncertainty is 1.9% and 2.0% for cycles 1 and 2 respectively.

While these 95% CI values fit within the same error levels observed in section 5, there are assemblies in both cycles where a linear fit is inadequate, as evidenced by high assembly RMS values as well as cases where multiple statistical tests are rejected in favor of an alternative hypothesis. The primary reason for such behavior is due to the volatile power history for both cycles that call into question the steady-state requirements for linear fitting to be appropriate. Moreover, a linear fit cannot be conducted without knowing the BEAVRS reaction rates a priori, and neglects taking into account any of the actual operating conditions within the reactor. Thus, these limitations

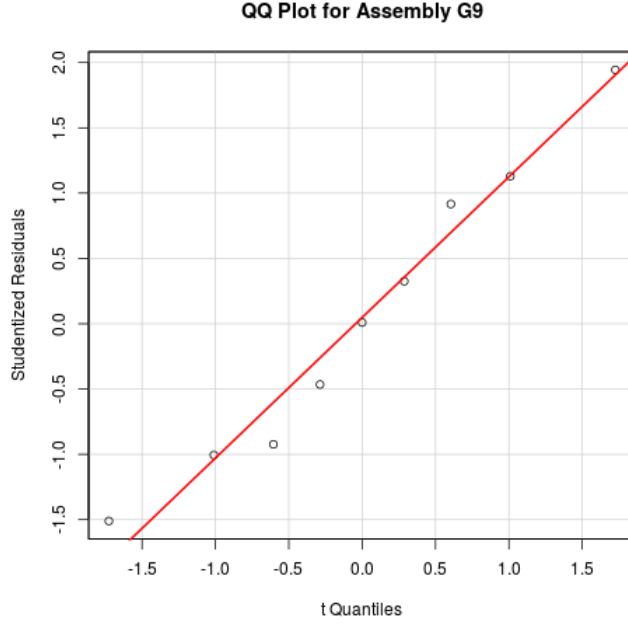


Figure 53: Q-Q plot for linear model applied to for assembly G9 in cycle 1.

necessitate the use of more predictive models that can be used to capture higher-order effects observed within the reactor. section 7.2 addresses such drawbacks by computing reaction rates from CASMO/Simulate to use as bases for comparison against BEAVRS data.

7.4 Summary of Results

This section serves to summarize the primary results from this task. Two models are used to compute detection uncertainty, namely the CASMO/Simulate model and the linear model. The 95% confidence estimates of the UQ methods based on the CASMO/Simulate model and linear regression model are presented in Table 22. Similarly, traditional methods for uncertainty quantification based on work from Section 5 are also presented in Table 21 for comparison. Theoretical analysis of axial uncertainties computes uncertainty using all measurements based on variations of signal strengths by groups. This approach determines uncertainties at each axial point using lookup tables and also sums each axial point and associated uncertainty to calculate axially integrated values. On the other hand, the multiple measurements method computes uncertainty for measurements based on repeated measurements that are conducted by the same detector in the same assembly at a given burnup.

It should be noted that results from the CASMO/Simulate model have been separated into two categories to show how much of the error stems from detection deviation and how much this error level increases after accounting for the uncertainty in using tilt-corrected data. The second column in Table 22 indicates the 95% confidence interval of $(\delta_\phi/\phi)_{ij,detection}$ in Equation 15, while the third column represents the 95% confidence interval of the cumulative time-dependent uncertainty,

p-Values for Various Statistical Tests			
Assembly	Breusch-Pagan Test	Durbin-Watson Test	Shapiro-Wilk Test
G9	0.560	0.826	0.885
G8	0.783	0.630	0.325
B8	0.341	0.974	0.878
B9	0.033	0.398	0.488
E9	0.125	0.094	0.567
E8	0.089	0.694	0.895
F10	0.579	0.250	0.372
D10	0.176	0.458	0.249
B12	0.217	0.104	0.565
B13	0.381	0.310	0.240
B10	0.127	0.186	0.655
C9	0.954	0.738	0.000
C8	0.295	0.014	0.201
F8	0.593	0.460	0.518
F9	0.178	0.484	0.792
E11	0.523	0.412	0.933
E10	0.166	0.274	0.973
C13	0.043	0.110	0.122
C12	0.488	0.572	0.259
C11	0.964	0.044	0.082
C10	0.429	0.256	0.883
D12	0.931	0.082	0.419
A11	0.107	0.780	0.473
A10	0.057	0.344	0.203
A9	0.659	0.784	0.167
A8	0.049	0.102	0.150
D8	0.760	0.952	0.218
D9	0.206	0.752	0.906

Table 20: Summary of p-values from statistical tests that test validity of linear regression model.

$(\delta_\phi/\phi)_{ij,time-dep}$, which accounts for the uncertainty in using tilt-corrected data as well.

Traditional Methods of Uncertainty Quantification		
	Theoretical Analysis of Axial Uncertainties	Multiple Measurements
Cycle 1	1.4%	1.8%
Cycle 2	1.4%	1.5%

Table 21: Summary of results from uncertainty quantification using traditional methods.

Uncertainty Quantification Methods Based on Time Series Analysis			
	CASMO/Simulate Model (before tilt-correction UQ)	CASMO/Simulate Model (after tilt-correction UQ)	Linear Regression Model (after tilt-correction UQ)
Cycle 1	0.8%	1.8%	1.9%
Cycle 2	0.9%	1.9%	2.0%

Table 22: Summary of results from uncertainty quantification using new methods based on time series analysis.

Tables 21 and 22 show that the uncertainty within the BEAVRS benchmark ranges from 1.4% to 1.8% in cycle 1 and 1.4% to 1.9% in cycle 2. These results indicate that time series analysis methods can be used to model time-dependent behavior, with measured data behaving in accordance with data from simulation tools. The CASMO/Simulate model, for example, predicts reaction rates based on operational conditions, and when accounting for the model bias, the time-dependent uncertainty produces uncertainty results on the same order as methods based on reaction rate map uncertainties. These findings are also independent of type of simulation model used to generate reaction rates. What is more, if uncertainties from using tilt-corrected data could be completely explained, then the remaining detection uncertainty from fitting the CASMO/Simulate model to BEAVRS data yield much lower errors than using traditional UQ methods.

References

- [1] Nicholas E. Horelik, Bryan Herman, Benoit Forget, and Kord Smith. Benchmark for evaluation and validation of reactor simulations. Technical report, Massachusetts Institute of Technology, 2013.
- [2] Nicholas E. Horelik, Bryan R. Herman, Benoit Forget, and Kord Smith. Benchmark for evaluation and validation of reactor simulations. In *International Conference on Mathematics and Computational Methods Applied to Nuclear Science & Engineering (M&C 2013)*, 2013.
- [3] S.G. Zimmerman, J.C. Brittingham, and M.L. Reed. Arizona public service company pwr reactor physics methodology using casmo4/simulate3. Technical report, 1999.
- [4] M Mitelman. In-core monitoring detectors. Technical report, 2001.
- [5] Jingang Liang, Shikhar Kumar, Benoit Forget, and Kord Smith. Integral full core multi-physics pwr benchmark with measured data beavrs uq neup 14-6742: Year 2 report. Technical report, Massachusetts Institute of Technology, 2016.
- [6] Shikhar Kumar, Jingang Liang, Benoit Forget, and Kord Smith. Integral full core multi-physics pwr benchmark with measured data beavrs uq neup 14-6742: Year 3 report. Technical report, Massachusetts Institute of Technology, 2017.
- [7] Jingang Liang, Koroush Shirvan, Miriam Rathbun, Matt Ellis, Hwansoo Lee, Benoit Forget, and Kord Smith. Integral full core multi-physics pwr benchmark with measured data beavrs uq neup 14-6742: Year 1 report. Technical report, Massachusetts Institute of Technology, 2015.
- [8] Elliot Sykora. Testing the EPRI Reactivity Depletion Decrement Uncertainty Methods. Master’s thesis, Massachusetts Institute of Technology, 2015.
- [9] Joel Rhodes, Kord Smith, and Deokjung Lee. Casmo-5 development and applications. In *PHYSOR 2006*, pages 10–14, 2006.
- [10] JT Cronin, KS Smith, DM Ver Planck, JA Umbarger, and M Edenius. Simulate-3 methodology, advanced three-dimensional two-group reactor analysis code. *Studsvik/SOA-95/18, Studsvik of America, Inc*, 1995.
- [11] Benjamin Collins. Personal Communication, 2017.
- [12] Benjamin Collins, T Downar, Jess Gehin, Andrew Godfrey, Aaron Graham, Daniel Jabaay, Blake Kelley, Kevin Clarno, Kang Seog Kim, Brendan Kochunas, et al. Mpact theory manual. Technical report, Technical Report CASL-U-2015-0078-000, Consortium for Advanced Simulation of LWRs, 2015.
- [13] M.J. Driscoll, T.J. Downar, and E.E. Pilat. *The Linear Reactivity Model for Nuclear Fuel Management*. American Nuclear Society, 2000.
- [14] Jonathan D. Cryer and Kung-Sik Chan. *Time Series Analysis With Applications in R*. Springer Science, 2008.
- [15] James D. Hamilton. *Time Series Analysis*. Princeton University Press, 1994.

- [16] Frank Harrell. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Springer, 2015.
- [17] Julian J Faraway. *Linear Models with R*. CRC press, 2014.
- [18] Trevor S Breusch and Adrian R Pagan. A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, pages 1287–1294, 1979.
- [19] R Dennis Cook and Sanford Weisberg. Diagnostics for heteroscedasticity in regression. *Biometrika*, 70(1):1–10, 1983.
- [20] James Durbin and Geoffrey S Watson. Testing for serial correlation in least squares regression. i. *Biometrika*, 37:409–428, 1950.
- [21] JP Royston. Algorithm as 181: the w test for normality. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 31(2):176–180, 1982.
- [22] Patrick Royston. Remark as r94: A remark on algorithm as 181: The w test for normality. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 44(4):547–551, 1995.