Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Evaluating Machine‐Learned Nuclear Data Precision in Full‐Core Nuclear Reactor Calculations: Computational Efficiency, Criticality and Burn‐up Analysis

Alexander Hashemi

alexander.hashemi@tum.de

Technical University of Munich

Rafael Macián-Juan

Technical University of Munich

Martin Ohlerich

Bavarian Academy of Sciences and Humanities

Article

Keywords: Nuclear Data, Machine Learning, OpenMC, Monte Carlo Transport, Full-Core Analysis, Pressurized Water Reactors

Posted Date: October 14th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7831183/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on January 10th, 2026. See the published version at https://doi.org/10.1038/s41598-026-35227-9.

Evaluating Machine-Learned Nuclear Data Precision in Full-Core Nuclear Reactor Calculations: Computational Efficiency, Criticality and Burn-up Analysis

Alexander Hashemi[1,*], **Rafael Macian-Juan**[1], **and Martin Ohlerich**[2]

[1]Chair of Nuclear Engineering, Technical University of Munich (TUM), Garching, Munich, D-85748, Germany
[2]Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, Garching, Munich, D-85748, Germany
[*]alexander.hashemi@tum.de

ABSTRACT

This study evaluates the novel machine learning based reduction of cross-sections and energy grid of continuous-energy nuclear data for one year full-core Monte Carlo criticality and burn-up analysis using OpenMC. The approach edits OpenMC's ENDF/B-VII.1 Hierarchical Data Format, version 5 (HDF5) nuclear data files, retaining ∼10% to 50% of nuclear data for 23 nuclides while preserving thresholds and resonances. EPR and VVER-1000 full-core models benchmark reduced nuclear data library against the original (windowed multipole disabled), to quantify performance and fidelity. Wall time decreased by 17.81% in EPR and 42.5% in VVER-1000. Peak memory (MaxRSS) decreased by 4.4% in EPR and increased by 5.0% in VVER-1000. The maximum absolute difference in $k_{\text{eff}}(t)$ for VVER-1000 remains within **96.79 pcm** at all times. VVER-1000 end of cycle reaction rates relative differences found for U-235 (n, f) 0.0017%, U-238 (n, f) 0.0605%, Xe-135 (n, γ) 0.0128%, Sm-149 (n, γ) 0.03%. Inventories EOC relative difference were 0.0039% U-235, 0.0003% U-238, 0.0135% Xe-135, 0.0341% Sm-149. The EOC relative difference for the Plutonium vector has been analyzed. Results prove that the developed reduction method accelerates full-core analysis, reduces MaxRSS while maintaining fidelity in neutronics studies.

Keywords: Nuclear Data, Machine Learning, OpenMC, Monte Carlo Transport, Full-Core Analysis, Pressurized Water Reactors

Introduction

High-fidelity reactor physics calculations hinge on the quality, consistency, and computational footprint of continuous-energy nuclear data. In Monte Carlo (MC) transport codes such as OpenMC[1], evaluated nuclear data are commonly stored in an HDF5 format[2] derived from A Compact ENDF (ACE) files generated via NJOY[3] processing and include tabulated reaction cross sections, fission neutron yields, secondary energy–angle distributions, and probability tables for the unresolved resonance range.[1,4] Performance at scale such as memory footprint and run time is affected not only by tally workload and particle statistics, but also by how cross sections are represented and searched during transport. OpenMC employs an energy grid acceleration known as *logarithmic mapping*: the incident neutron energy domain is split into equal lethargy segments, default 8,000 bins as recommended by Brown [5], and each segment is mapped to index bounds on every nuclide-specific grid to constrain per-collision lookups.[6] In addition to tabulated pointwise data, Windowed multipole (WMP) method [7] may further reduce memory and enable on-the-fly Doppler broadening in the resolved resonance region [6]; outside that region, temperature treatment reverts to pointwise tabulations or interpolation tables. [6] These algorithmic choices determine how much of the memory and run time is attributable to nuclear data access rather than to random-walk physics.

A recent study introduced an adaptive resampling, energy grid, and cross-sections thinning and direct HDF5 editing workflow for OpenMC libraries, demonstrating that intelligently pruning the energy grid can preserve resonance fidelity while markedly reducing memory and processing time.[8] Concretely, the method performs region-aware pruning of the pointwise energy grid with slope-based refinement near thresholds and resonances, then writes the reduced arrays directly back to OpenMC-compatible HDF5 without reusing an ACE. The full algorithm and validation appear in the companion paper and its supplementary material. [8] This prior work quantified interpolation errors on single-nuclide reactions and established significant savings relative to the legacy NJOY → ACE → HDF5 pipeline. The workflow diagram of this adaptive resampling, grid-thinning, and direct HDF5 editing algorithm is given in figure 1.

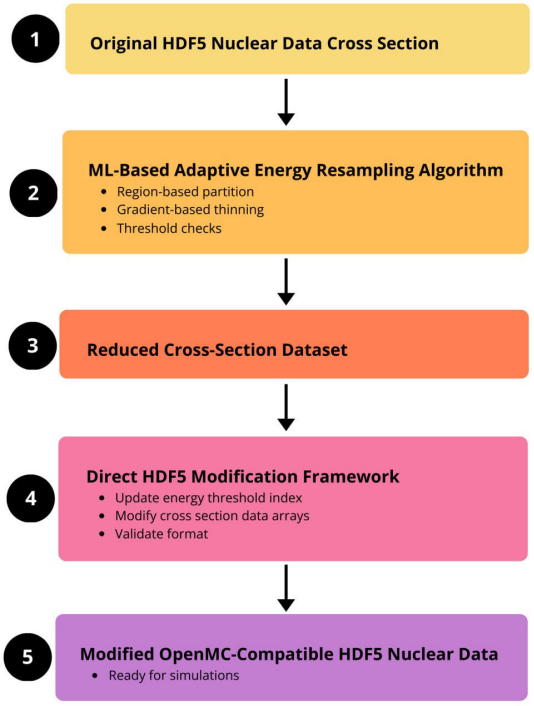


Figure 1. Simplified process flow diagram for Nuclear Data modification in HDF5 format.
[8]

The present paper evaluates the *physics and performance consequences* of using such machine-learned, pointwise-reduced nuclear data in *full-core* Monte Carlo depletion. Two full-core pressurized-water reactors (PWR) are used as references: a European Pressurized Reactor (EPR) and a VVER-1000 full-core model. The OpenMC input models supplied for this study implement these cores with standard materials and core layouts (illustrative plots provided in figure 2 and figure 3). The continuous-energy library is ENDF/B-VII.1 as distributed for OpenMC. The main question is whether retaining roughly 10–50% of energy-grid and cross-sections for selected 23 nuclides can reduce wall-clock time and peak memory (MaxRSS) while preserving integral measures multiplication factor $k_{\text{eff}}(t)$ and isotopic inventories over a one-year irradiation burn-up.
This investigation quantifies (i) changes in wall-clock time and peak resident memory (MaxRSS) and (ii) deviations in $k_{\text{eff}}(t)$ and nuclide inventories over depletion when replacing the original (unmodified) ENDF/B-VII.1 HDF5 library with a modified library in which only the energy grids and tabulated cross sections of a targeted nuclide set are thinned. Geometry, tallies, source settings, and all simulation controls are held fixed between the baseline and reduced-data cases. The study does *not* re-evaluate ENDF data, re-process NJOY tolerances, propagate covariance/uncertainty data, compare alternative acceleration features (e.g., WMP), or attempt variance-reduction/solver changes; such topics are outside the present scope.
In a continuous-energy eigenvalue calculation, reaction sampling at each collision draws on the microscopic cross sections available for each nuclide and reaction channel (e.g., elastic, (n,γ) capture, (n,f) fission, inelastic, $(n,2n)$), with fissionable nuclides carrying $\nu(E)$, prompt/delayed components, and secondary spectra from the nuclear data library or from a multipole representation where enabled. At a collision, reaction x is sampled with probability $P(x) = \sigma_x(E)/\sigma_t(E)$ using the microscopic values at the sampled energy. The total cross section decomposes as $\sigma_t(E) = \sigma_s(E) + \sigma_a(E)$, where *Absorption* includes fission in addition to non-scattering disappearance channels (e.g., capture, (n,α))).[6,9].
For burn-up, OpenMC tallies the transmutation reaction rates that appear in the depletion chain[10] and couples them to an ODE integrator (e.g. CRAM/predictor–corrector families, predictor, CE/CM, EPCRK4, LE/QI) to evolve number densities.[11,12] Hence, the fidelity of (n,f), $\nu\sigma_f$, and (n,γ) on the major actinides, as well as of the strong absorbers among fission products, dominates impacts on k_{eff} and inventories. In thermal PWRs, short-lived Xe-135 and the longer-lived Sm-149 are well-known parasitic absorbers with large thermal capture cross sections, shaping short- and medium-term reactivity, respectively.[13]

Methods

Reference nuclear data and reduction strategy

All calculations use the official OpenMC version 15.2 with ENDF/B-VII.1 incident-neutron HDF5 library, which provides pointwise cross sections and tabulated temperatures for each nuclide, along with secondary distributions and, where applicable,

URR probability tables and $S(\alpha,\beta)$ kernels for bound scatterers.[1,4] The reduced library is obtained by applying the adaptive resampling procedure introduced in the companion paper [8] to the *per-nuclide* energy grid and tabulated reaction cross sections; thresholds and resonance structures are retained through region-weighted, gradient-based selection.[8] The resulting HDF5 files remain OpenMC-compatible in which only the designated nuclides are thinned; all dataset attributes required by the OpenMC readers remain intact.

To isolate the effect of pointwise reduction, the windowed multipole (WMP) data[6,14] is disabled in transport. In OpenMC, the windowed-multipole capability is controlled by the settings entry `temperature['multipole']`; if set to `True` and multipole data are available, OpenMC evaluates resolved-resonance cross sections on the fly, otherwise standard pointwise tabulations are used. The energy grid search uses the default logarithmic mapping with `n_log_bins = 8000`.[6] In this study, the WMP is set to `False`, and the temperature method follows pointwise *interpolation/nearest* treatment as in standard use, so that any performance change is attributable solely to grid/cross-section thinning.

Edited nuclide set

The edited nuclides set comprises the principal fissile/fertile actinides and the most influential absorbers that control PWR reactivity over operating and depletion time scales:

^{233}U, ^{234}U, ^{235}U, ^{236}U, ^{238}U, ^{238}Pu, ^{239}Pu, ^{240}Pu, ^{241}Pu, ^{242}Pu,^{135}Xe, ^{149}Sm,

^{131}I, ^{103}Rh, ^{143}Nd, ^{145}Nd, ^{133}Cs, ^{149}Pm, ^{157}Gd,^{237}Np, ^{241}Am, ^{243}Am, ^{244}Cm

Major actinides dominate k_{eff} via (n,f) and (n,γ); U-238$(n,\gamma)\rightarrow$ U-239 $\rightarrow$ Np-239$\rightarrow$ Pu-239 links fertile capture to plutonium breeding, altering spectrum and reactivity. Fission-product poisons Xe-135 and Sm-149 set short- and medium-term reactivity swing; burnable absorbers (e.g., Gd-157) and lanthanides (Nd, Pm, Sm) shape spectrum and long-term worth. Minor actinides (Np, Am, Cm) affect capture sinks and decay heat, with inventories that are sensitive to capture branching. This set thus captures the primary sensitivities in thermal PWR cores while limiting the scope of HDF5 edits to species that materially impact keff and inventories. Authoritative reactor-physics treatments identify Xe-135 and Sm-149 as the dominant fission-product poisons in thermal systems; the remaining selections reflect capture/fission importance in standard depletion chains.[11,13]

Amount of reduction and rationale

Pointwise reduction targets a retained-grid fraction of approximately 10–50% per affected nuclide and temperature, implemented as a fixed point budget on each nuclide's energy grid. In practice, a retained count near 6,000 energies per temperature provided a robust compromise across heavy-nuclide resonance structure and smooth fast-range behavior in preliminary trials, delivering significant compression without detectable bias in single-nuclide verification.[8] Coarser targets ($\lesssim$3,000) risk resonance under-sampling in U/Pu isotopes; finer targets ($\gtrsim$10,000) diminish the performance benefit. The first paper reported that, for individual reactions, adaptive thinning to $\sim$1,500 points at a given temperature preserved high-resolution features with small regression errors relative to the originals; the present work extends that result to full-nuclide energy grids and full-core depletion. Table 1 shows the approximate reduction of 23 nuclides used, tabulated energy grid points per nuclide in the OpenMC ENDF/B-VII.1 library in this study.

Nuclide	N_{orig}	N_{red}	ΔN	Reduced [%]	Retained [%]
^{133}Cs	49,377	6,000	43,377	87.8	12.2
^{103}Rh	56,237	6,000	50,237	89.3	10.7
^{149}Pm	2,465	1,200	1,265	51.3	48.7
^{145}Nd	42,582	6,000	36,582	85.9	14.1
^{143}Nd	35,916	6,000	29,916	83.3	16.7
^{131}I	2,557	1,200	1,357	53.1	46.9
^{244}Cm	28,851	6,000	22,851	79.2	20.8
^{243}Am	51,061	6,000	45,061	88.2	11.8
^{241}Am	37,008	6,000	31,008	83.8	16.2
^{237}Np	120,834	6,000	114,834	95.0	5.0
^{157}Gd	16,180	4,000	12,180	75.3	24.7
^{135}Xe	2,630	2,000	630	24.0	76.0
^{149}Sm	41,075	14,000	27,075	65.9	34.1
^{233}U	45,002	6,000	39,002	86.7	13.3
^{234}U	45,977	6,000	39,977	86.9	13.1
^{235}U	60,631	24,000	36,631	60.4	39.6
^{236}U	41,854	6,000	35,854	85.7	14.3
^{238}U	133,950	67,000	66,950	50.0	50.0
^{238}Pu	13,470	6,000	7,470	55.5	44.5
^{239}Pu	59,832	24,000	35,832	59.9	40.1
^{240}Pu	45,096	6,000	39,096	86.7	13.3
^{241}Pu	15,625	7,800	7,825	50.01	49.9
^{242}Pu	14,648	6,000	8,648	59.0	41.0

Table 1. Approximate reduction of tabulated energy-grid points per nuclide for 23 nuclides in the OpenMC ENDF/B-VII.1 library used in this study. The column N_{orig} is the original number of tabulated energies, N_{red} is the size after applying the adaptive resampling and direct HDF5 edit, $\Delta N = N_{\text{orig}} - N_{\text{red}}$, and the last two columns report the percentage reduced and retained.

OpenMC depletion solver

OpenMC's depletion module assembles the requisite reaction rates from transport (for all reactions present in the depletion chain and available in the library) and advances nuclide number densities using matrix exponential methods; predictor and higher-order integrators are available.[12] The models here use one-year irradiation with fixed power normalization and standard step partitioning (monthly cadence after any initial short steps), identical between baseline and reduced-data cases. Reaction pathways are taken from the standard ENDF/B-VII.1 depletion chain for thermal spectrum.

Implementation Details: OpenMC Full-Core models and inputs

In this study, two full-core OpenMC inputs with fresh fuels are employed: **EPR full core** and **VVER-1000 full core**. OpenMC model input of an EPR is given in table 2. OpenMC input model of a VVER-1000 summarized in table 3. Both inputs are run twice with identical geometry, source, batches, and depletion schedules, differing only in the path to the nuclear data library: (i) original ENDF/B-VII.1 HDF5 (baseline) and (ii) modified HDF5 with reduced energy grids/cross sections for the 23-nuclide set. To avoid tally-memory confounding of MaxRSS comparisons, *both* production runs (EPR and VVER-1000) were executed with all diagnostic tallies disabled. The VVER-1000 input contains optional tallies that are commented out in production. The plot for OpenMC model of an EPR is given in figure 2. Also, a plot for OpenMC model of VVER-1000 is shown in figure 3.

Simulation matrix and metrics

For each core (EPR, VVER-1000) two campaigns are executed:

1. **Baseline:** Original, unmodified ENDF/B-VII.1 HDF5; no WMP; `n_log_bins`=8,000; depletion over one year.
2. **Reduced sized nuclear data:** Identical input setup , but with edited nuclear data for the 23 nuclide set retains ∼10–50% of the original energy grid points, with reaction tables re-tabulated at the reduced grid.

Performance is quantified by wall-clock time and peak resident memory (MaxRSS) recorded from the batch scheduler logs; *physics* is assessed by $k_{\text{eff}}(t)$ and nuclide inventories $N_i(t)$ for key actinides and poisons. All random seeds, depletion step sizes, source particle counts, inactive/active batch counts, and normalization settings are held fixed across the two campaigns per core. To attribute any time/memory change *uniquely* to cross-section/energy grid thinning, all such features are left at their documented defaults (logarithmic mapping active; WMP=False).

Category	Specification
Reactor / core layout	Simplified EPR PWR (OECD/NEA Monte Carlo Performance benchmark) with **241** 17×17 fuel assemblies. Square fuel-pin pitch **1.26 cm**; assembly pitch **21.42 cm**. Active core height **420 cm** ($z = \pm 210$ cm). Cylindrical radii (cm): core-lattice limit $r = 209$, inner/outer RPV $r = 229/249$; additional internal boundary at $r = 187.6$ (reflector interface). Axial planes: vessel bottom $z = -250$, bottom nozzle -242, bottom skirt -230, active bottom -210, core midplane 0, active top 210, top skirt 230, top nozzle 242, vessel top 250 (vacuum at ± 250).
Fuel pins	UOX fuel, 4.95 wt% ^{235}U, $\rho = 10.5$ g/cm^3, $T = 900$ K; pellet radius $r = 0.4095$ cm. Zircaloy cladding, $\rho = 5.77$ g/cm^3, $T = 613.15$ K; cladding outer radius $r = 0.475$ cm. Each 17×17 assembly contains **265** fuel rods and **24** guide tubes; no instrument tube in this input. Guide-tube radii: $r = 0.56/0.62$ cm (inner/outer).
Coolant / moderator	Cold borated water (inlet): number density 0.07416 atom/b-cm, $T = 560$ K; Hot borated water (outlet): 0.06614 atom/b-cm, $T = 590$ K; both include B-10 and B-11 and S(α, β) for H in H_2O.
Reflectors / structures	Lower radial reflector ($\rho = 4.32$ g/cm^3, $T = 560$ K), upper radial/top-plate region ($\rho = 4.28$ g/cm^3, $T = 590$ K), bottom plate ($\rho = 7.184$ g/cm^3, $T = 560$ K), bottom nozzle ($\rho = 2.53$ g/cm^3, $T = 560$ K), top of fuel assemblies ($\rho = 3.044$ g/cm^3, $T = 590$ K), top nozzle ($\rho = 1.746$ g/cm^3, $T = 590$ K), reactor pressure-vessel steel ($\rho = 7.9$ g/cm^3, $T = 575$ K).
Monte Carlo settings	**Batches** = 100, **inactive** = 50, **particles/batch** = 200,000; source: uniform box $[-160, -160, -183]$ to $[160, 160, 183]$, `only_fissionable=True`. Temperature treatment: `multipole=False` (WMP disabled), `method='interpolation'`, range 290–1500 K.
Nuclear data / depletion	ENDF/B-VII.1 HDF5 library; depletion chain `chain_endfb71_12_pwr.xml` for thermal spectrum; constant thermal power **4.59 GW$_{th}$**; Predictor integrator; time grid: 30 daily steps + 11 monthly steps (≈1 year, 41 steps). Fuel volume set as $265 \times \pi r^2 h$ per assembly, scaled to 241 assemblies with $r = 0.4095$ cm, $h = 420$ cm.
Tallies / plots	Tallies disabled for production. XY mesh tally (100×100) for flux and fission over the mid-plane; core spectrum tally with 500 log-spaced groups (1e-5 eV–20 MeV); material-colored geometry plots exported for QA.

Table 2. OpenMC EPR PWR full-core model: geometry, materials and simulation input.[15]

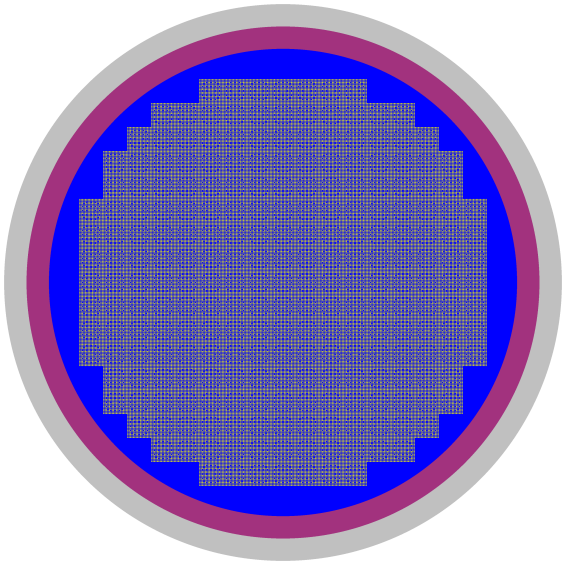


Figure 2. EPR full-core.

Category	Specification
Reactor / core layout	VVER-1000, full-core hexagonal lattice with **163** fuel assemblies; hex assembly pitch **23.6 cm**; active core height **353 cm** (planes $z = \pm 176.5$ cm); geometric core diameter **340 cm**; stainless-steel radial reflector thickness **23.6 cm**; vacuum radial boundary at $r = 170 + 23.6 = 193.6$ cm.
Fuel rods (depletable)	Material `au13` (UO_2): $\rho = 10.4$ g/cm^3, $T = 900$ K; weight fractions U-235 3.87%, U-238 84.60%, O-16 11.53%; annular fuel pellet with $r_{in} = 0.075$ cm and $r_{out} = 0.3775$ cm; He gap; Zr-Nb-Hf alloy cladding with $r_{in} = 0.386$ cm and $r_{out} = 0.455$ cm.
Burnable/absorber rods	Material `av5` (Al/B/Cr/Fe/Ni/Zr alloy): $\rho = 2.6989$ g/cm^3, $T = 600$ K; B-10 = 0.2 wt%, B-11 = 1.0 wt%; distributed in specified rings; central instrumentation tube present.
Coolant / moderator	Light water: $\rho = 0.727$ g/cm^3, $T = 600$ K; S(α,β) treatment for H in H_2O enabled.
Structural / reflector	SS304 reflector ($\rho = 7.90$ g/cm^3, $T = 600$ K); cladding alloy $\rho = 6.452$ g/cm^3, $T = 600$ K.
Pin/assembly statistics	Hex assembly with **330** rods + **1** central tube (asserted in input); intra-assembly pin pitch **1.275 cm**.
Monte Carlo settings	**Batches** = 80, **inactive** = 20, **particles/batch** = 250,000; source: uniform box $[-170,-170,-176.5]$ to $[170,170,176.5]$, `only_fissionable=True`. Temperature treatment: `multipole=False` (WMP disabled), `method='nearest'`, tolerance 10 K, range 290–1500 K.
Nuclear data / depletion	ENDF/B-VII.1 HDF5 library; depletion chain `chain_endfb71_12_pwr.xml`; constant thermal power **3.0 GW$_{th}$**; Predictor integrator; time grid: 30 daily steps + 11 monthly steps (≈1 year, 41 steps). Fuel volume set from pin geometry and number of assemblies.
Tallies / plots	Tallies disabled for production depletion (blocks commented); assembly and full-core XY/YZ plots exported for QC.

Table 3. OpenMC VVER-1000 full-core model: geometry, materials and simulation input.[16]

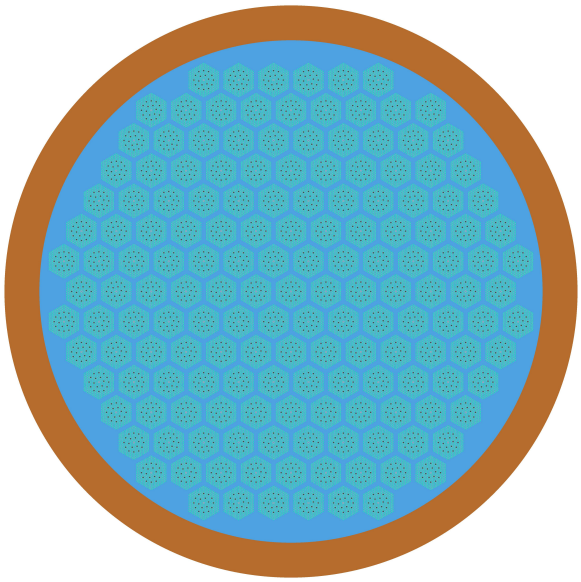


Figure 3. VVER-1000 full-core.

Assumptions, validity domain, and potential failure modes

Tables 2 and 3 provide details of EPR and VVER-1000 full-core inputs in OpenMC version 15.2 .

Assumptions.

- **Nuclear data and transport representation.** All runs use OpenMC's version 15.2 ENDF/B-VII.1 HDF5 data; WMP is disabled to isolate the effect of thinning pointwise data. energy grid search uses the default equal-lethargy mapping (`n_log_bins` not altered in the inputs). The only change between the "baseline" and "reduced-data" campaigns is the path to the nuclear data folder (original vs. edited HDF5 files).

- **Geometry and materials.** *EPR:* a 21×21 assembly grid with water assemblies at the radial periphery; axial modeling uses hot/cold halves; outer radial/axial boundaries are vacuum. Fuel is UOX at 4.95 wt% with Zircaloy cladding; borated water contains ^{10}B/^{11}B and S(α,β) for H in H_2O. A fixed fuel volume is assigned as $265 \times \pi r^2 h \times 241$ using $r = 0.4095$ cm and $h \approx 420$ cm. *VVER:* a hexagonal lattice with a single instrument tube and annular fuel; a stainless-steel cylindrical reflector surrounds the core; top/bottom boundaries are vacuum. Pin radii ($r_{\text{in}} = 0.075$ cm, $r_{\text{out}} = 0.3775$ cm) and active height $h = 353$ cm set the depletable fuel volume; assembly and core counts are computed explicitly in the input.

- **Depletion model.** One-year depletion uses `openmc.deplete.PredictorIntegrator` and `chain_endfb71_12_pwr.xml`; only fuel is marked depletable. The source distribution is uniform over the core bounding box and restricted to fissionable regions.

- **Instrumentation (tallies).** Optional diagnostic tallies (e.g., XY-mesh flux/fission and a 500-group spectrum tally) appear in the EPR input but are deactivated in output; production timing/memory comparisons are performed with no tallies recorded to avoid biasing the performance metrics.

- **Cross-section reduction target set.** Editing of the HDF5 libraries is applied only to the stated 23 nuclides (U-233, U-234, U-235, U-236, U-238, Pu-238, Pu-239, Pu-240, Pu-241, Pu-242, Xe-135, I-131, Sm-149, Rh-103, Nd-143, Nd-145, Cs-133, Pm-149, Gd-157, Np-237, Am-241, Am-243, Cm-244); all other nuclides remain unchanged. The reduction level preserves roughly 10-50% of the original energy grid/cross-section points according to the adaptive resampling method described previously.

Validity domain.

- **Reactor class and operating state.** Results are valid for large PWR-type thermal reactors under nominal, steady-state conditions with fixed soluble boron and without control-rod insertion or xenon/thermal-hydraulic transients. The EPR model is the simplified OECD/NEA benchmark configuration; the VVER model is a simplified VVER-1000 hex core.

- **Library temperatures and interpolation.** Accuracy requires that material temperatures lie within the tabulated temperature range of the ENDF/B-VII.1 HDF5 files. The EPR case uses linear interpolation, while the VVER case uses nearest-neighbor selection with a 10 K tolerance; conclusions on the effect of grid thinning strictly apply under those interpolation policies.

- **Depletion chain coverage.** The `chain_endfb71_12_pwr.xml` chain covers major actinides and important fission products used in power PWR analyses. Quantities relying on nuclides absent from the chain, or on high-order reaction channels not included, fall outside scope.

- **Statistical precision.** Monte Carlo uncertainties are controlled by the stated batch/particle settings; quantitative comparisons of k_{eff} and inventories are intended at the level of the reported one-sigma confidence intervals. Reproducibility at the bit-level requires fixing the random seed; otherwise, runs are statistically but not bitwise reproducible.

- **Objective of the study.** The inputs are designed to isolate performance and physics differences arising from *nuclear-data thinning only*. Other acceleration features (e.g., multipole on-the-fly broadening, unionized grids) are intentionally disabled or left at defaults.

Potential failure modes and sources of bias.

- **Temperature–library mismatch.** If material temperatures fall outside the HDF5 tabulation range, OpenMC may extrapolate or snap to the nearest temperature (VVER input), introducing bias unrelated to nuclear-data thinning. This risk grows if coolant/fuel temperatures are later modified without regenerating consistent HDF5 datasets.

- **Fuel-volume specification.** Depletion densities scale with the user-assigned material volume. In the EPR deck, the rod count per assembly is fixed at 265 for this simplified model; any deviation from the actual pin count of a target reference would bias atom densities and reaction-rate normalizations. In the VVER deck, assembly pin counts are programmatically computed; changing ring patterns without updating the volume logic will propagate to all inventory results.

- **Geometry idealizations.** The simplified EPR and VVER cores omit explicit control rods, burnable absorbers, detailed reflector heterogeneities, and core internals. If the reduced-data library is later applied to problems where these features materially affect spectra (e.g., strong absorber insertion), the present error bounds may not hold.
- **Tally overhead.** Enabling diagnostic tallies (mesh or energy-spectrum) increases memory and wall time. For fair performance comparisons between original and reduced libraries, all tallies must remain disabled; otherwise, differences in tally data structures could mask benefits from data thinning.
- **Cross-section editing scope.** Only the 23 designated nuclides are thinned. If other nuclides later become rate-limiting for transport (e.g., reflector isotopes, coolant boron under unusual borations), the observed savings and accuracy may change.
- **Chain coverage and spectral shifts.** The chain file governs available transmutation pathways. Strong spectral shifts (e.g., deep control-rod insertion, spectral hardening by heavy poisons) or long campaigns beyond one year could increase the importance of nuclides/reactions not explicitly represented, weakening conclusions on inventory errors.
- **Random-number control.** Absence of a fixed seed can yield small run-to-run differences in k_{eff} and inventories. If seeds differ between the baseline and reduced-data runs, apparent biases may reflect statistical scatter rather than nuclear-data edits.
- **Algorithmic interaction with OpenMC internals.** Because windowed-multipole is disabled, all conclusions pertain to pointwise transport. If WMP is re-enabled later, OpenMC will represent the resolved-resonance region differently; the performance/accuracy envelope of the reduced libraries would need re-evaluation in that regime.

Reproducibility and traceability to the developed method algorithm[8]

The machine learning resampling algorithm, the HDF5 modification framework and OpenMC version 15.2 inputs were implemented in Python 3.9, leveraging libraries such as NumPy [17] and SciPy [18] for numerical computations and interpolation, h5py [19] for direct interaction with HDF5-based nuclear data, Scikit-Learn [20] for machine learning model training and benchmarking, and Matplotlib [21] for visualization of resampled cross-sections and error metrics. The implementation was thoroughly tested on a high-performance computing cluster at the Leibniz Supercomputing Centre, ensuring its scalability for large-scale nuclear data processing. All input files for the EPR and VVER-1000 cases, along with the author's paper for the developed algorithm implementing the HDF5 reduction[8], are archived as part of this project's materials and identified by filename in the text above. The methodology paper provides algorithmic details, hyper-parameters, and environment specifications sufficient to reproduce the reduced libraries and verify the nuclide benchmarks. The above assumptions and limits align with the adaptive resampling and direct HDF5-editing workflow reported previously (machine-learned, region-aware thinning of pointwise cross sections). Where this study differs is the application to full-core, one-year depletion with identical inputs aside from the HDF5 path, enabling clean attribution of time/memory and physics differences to nuclear-data reduction alone.

Results and Discussion

This study assesses the physics impact of a developed method, an adaptive resampling algorithm, and a direct HDF5 modification framework[8] to reduce pointwise nuclear data libraries on the full-core criticality and burn-up analysis for the VVER-1000 and EPR over one year. In the main paper, the analysis over 365 days for the VVER-1000 is provided. The analysis of EPR and VVER-1000 cases with a much radical energy grid and cross-section reduction results is provided in the supplementary information. The evaluations concentrate on (i) computational performance studies by measuring the total processing time and peak memory (RSS), (ii) the time evolution of the neutron multiplication factor $k_{\text{eff}}(t)$, (iii) core integrated reaction rates for the channels most critical to reactivity and burn-up such as fission rate ((n, f)) of U-235 and U-238, as well as capture rate ((n, γ)) of Xe-135 and Sm-149, and finally (iv) isotopic inventories for the nuclides U-235, U-238, Pu-238, Pu-239, Pu-240, Pu-241 and fission-product poisons Xe-135 and Sm-149 over one year.

Computational Performance

Table 4 shows that efficiency and computational performance were also evaluated by measuring the total processing time and peak memory (RSS) for both the VVER-1000 and EPR. All runs were executed on a single node Intel(R) Xeon(R) Platinum 8380 CPU (Ice Lake), `AllocCPUS`=112, 2×40 cores, 1TB RAM with 32 threads and no GPU acceleration. The results show that the proposed method of adaptive resampling and using reduced nuclear data, resulting in speed-ups by factors of **1.217**(reducing 17.81% wall-time) in EPR and **1.741** (reducing 42.5% wall-time) in VVER-1000, relative to using the original nuclear library under the same configuration (table4), while maintaining the same level of accuracy within all evaluated metrics.

Reactor	Library	Elapsed (h:m:s)	Wall [s]	MaxRSS [GiB]	Speed-up (×)	ΔWall [%] / ΔMaxRSS [%]
EPR	Original	18:17:39	65859	1.773	1.217	17.81% ↓ / 4.40% ↓
	Reduced	15:02:07	54127	1.695		
VVER-1000	Original	15:25:15	55515	2.722	1.741	42.5% ↓ / 5.0% ↑
	Reduced	08:51:47	31907	2.859		

Table 4. Wall-clock time and peak memory (MaxRSS) from `sacct` for full-core over 365 days depletion time with OpenMC using ENDF/B-VII.1. The "Reduced" library uses the ML-based Energy grid and cross-section thinning [8]. MaxRSS values converted from `K` units and shown in GiB.

Error metrics and units

The figures present results obtained using original nuclear data, reduced nuclear data, and their comparison, accompanied by pointwise error panels. For any time-dependent observable $x(t)$ (e.g. a reaction rate, or an isotope concentration number density), the pointwise relative absolute difference in percent is defined as

$$\mathrm{RelErr}(t) \;=\; 100 \times \frac{|x_{\mathrm{reduced}}(t) - x_{\mathrm{original}}(t)|}{|x_{\mathrm{original}}(t)|} \;\%.$$

Summary statistics reported are the mean over $t \in [0, 365]$ d, the 95th percentile, the maximum, and the end of cycle (EOC) value. For the effective multiplication factor k_{eff}, the *absolute* difference is reported in per-cent-mille (pcm), where $1\,\mathrm{pcm} = 10^{-5}$ in $\Delta k/k$.

$$|\Delta k|_{\mathrm{pcm}}(t) \;=\; 10^5 \left|k_{\mathrm{reduced}}(t) - k_{\mathrm{original}}(t)\right|.$$

Effective Multiplication Factor $k_{\mathrm{eff}}(t)$

Figure 4 shows the evolution of k_{eff} over 365 days for the VVER-1000. The time history of $|\Delta k|_{\mathrm{pcm}}(t)$ appears in the bottom-right panel of figure 4; summary statistics are given in table 5. Across the year, the reduced library is consistently lower by tens of pcm. Quantitatively in table 5, the mean $|\Delta k|$ is **29.21 pcm**, the 95th-percentile is **55.36 pcm**, the maximum is **96.79 pcm** at day **12**, and the EOC value is **11.12 pcm**. The persistent down-shift is consistent with a small decrease in net fission probability (or a slight increase in net absorption) introduced by grid thinning, while resonance-dominated regions are preserved. The pcm envelope in table 5 characterizes the scale of deviations to be expected in the governing reaction rates and inventories; it is not a strict bound.

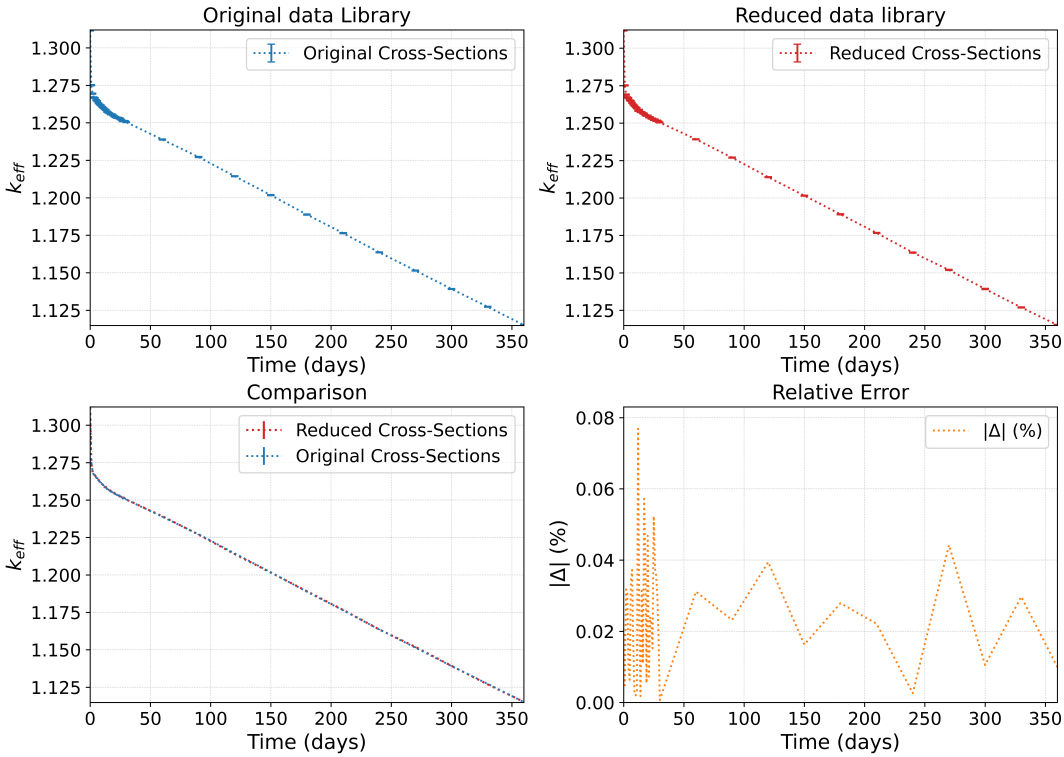


Figure 4. VVER-1000: Time evolution of k_{eff} for original vs. reduced nuclear data libraries over 365 days depletion time.

Observable	Mean $\lvert\Delta k\rvert$ [pcm]	95th [pcm]	Max $\lvert\Delta k\rvert$ [pcm] at day	EOC $\lvert\Delta k\rvert$ [pcm]
k_{eff}	29.21	55.36	96.79 at 12	11.12

Table 5. VVER-1000: Absolute difference for $k_{\text{eff}}(t)$ over 365 days in pcm.

Fission and Capture Reaction Rates

Four reaction rates that control reactivity in thermal PWRs are considered as fission rate(n,f) of U-235 and U-238, as well as capture rate (n,γ) of Xe-135 and Sm-149. Figures 5, 6, 7, 8 show the original vs. reduced time histories. Across channels, temporal means range from **0.0036%** (U-235 (n,f)) to **0.0556%** (U-238 (n,f)); the 95th-percentile absolute relative differences are $\leq$ **0.0954%** and the maxima are $\leq$ **0.1790%**. Maxima occur at **day 300** for U-235 (n,f), **day 28** for U-238 (n,f), **day 29** for Xe-135 (n,γ), and **day 210** for Sm-149 (n,γ). Table 6 shows the reaction rates statistics.

Isotope	Reaction	Mean [%]	95th [%]	Max [%]	Day of max	EOC [%]
U-235	(n,f)	0.0036	0.0079	0.0150	300	0.0017
U-238	(n,f)	0.0556	0.0954	0.1790	28	0.0605
Xe-135	(n,γ)	0.0232	0.0330	0.0656	29	0.0128
Sm-149	(n,γ)	0.0299	0.0436	0.0763	210	0.0300

Table 6. VVER-1000: Relative errors for reaction rates over 365 days.

The sub-percent temporal means and tight 95th-percentile bounds in table 6 indicate that the reduced pointwise representation preserves the flux lethargy weighting that controls integral rates across thermal, epithermal, and fast ranges. The timing of the maxima is reactor-physically consistent: an **early** peak for **Xe-135** (n,γ) at **day 29** reflects the initial I-135→Xe-135 buildup; a **BOL** maximum for **U-238** (n,f) at **day 28** tracks the early fast-fission contribution; a **mid-cycle** maximum for **Sm-149** (n,γ) at **day 210** corresponds to Pm-149 (β^-) feeding and absorber saturation; and **U-235** (n,f) peaks **late** at **day 300**. The magnitudes (means **0.0036–0.0556%**, 95th-percentiles $\leq$ **0.0954%**) remain well within the sub-percent envelope established by the k_{eff} analysis.

U-235 (n,f) **fission rate**: Relative differences are extremely small and peak late in the cycle: *mean* **0.0036%**, *95th* **0.0079%**, *maximum* **0.0150%** at **day 300**, *EOC* **0.0017%**. The late, low-magnitude maximum is consistent with gradual spectral hardening and cumulative fuel evolution.

U-238 (n,f) **fission rate**: Differences peak early and then diminish: *mean* **0.0556%**, *95th* **0.0954%**, *maximum* **0.1790%** at **day 28**, *EOC* **0.0605%**. The early maximum reflects BOL fast-fission leverage before plutonium growth and spectrum evolution reduce its relative role.

Xe-135 (n,γ) **capture rate**: Deviations remain modest and peak early: *mean* **0.0232%**, *95th* **0.0330%**, *maximum* **0.0656%** at **day 29**, *EOC* **0.0128%**. The early maximum follows the I-135→Xe-135 transient; the small EOC value indicates preserved thermal capture behavior.

Sm-149 (n,γ) **capture rate**: Differences peak at mid-cycle and remain small: *mean* **0.0299%**, *95th* **0.0436%**, *maximum* **0.0763%** at **day 210**, *EOC* **0.0300%**. This timing matches slower Pm-149 feeding and absorber saturation.

The small rate deviations are consistent with the reduction workflow, which preserves reaction thresholds and concentrates retained points where cross sections vary most rapidly (resonance slopes) while thinning aggressively in smooth fast/continuum ranges. As a result, interpolation error is minimized precisely in the energy bands that dominate the lethargy-weighted flux for (n,f) on U isotopes and (n,γ) on strong absorbers. Because OpenMC's equal-lethargy mapping constrains per-collision lookups to local index bounds, keeping dense support across resolved resonances ensures that the local ratio $\sigma_x(E)/\sigma_t(E)$ is stable against thinning, which in turn stabilizes integral rates over the cycle.

The sub-percent means and $< 0.18\%$ maxima also align with the observed $|\Delta k|_{\text{pcm}}$ envelope: time windows with slightly larger relative differences coincide with periods when the underlying rate is evolving rapidly (e.g. early xenon build-in, BOL prominence of U-238 fast fission, mid-cycle Pm-149→Sm-149 feed-in). Even in those windows, the rate perturbations remain small, indicating that grid thinning re-weights reaction competition only marginally. Within the present (deterministic) comparison—no per-step MC error bars recorded—the rate differences are therefore best interpreted as controlled, algorithm-induced perturbations that are reactor-physically consistent with the tens of pcm spread in k_{eff}.

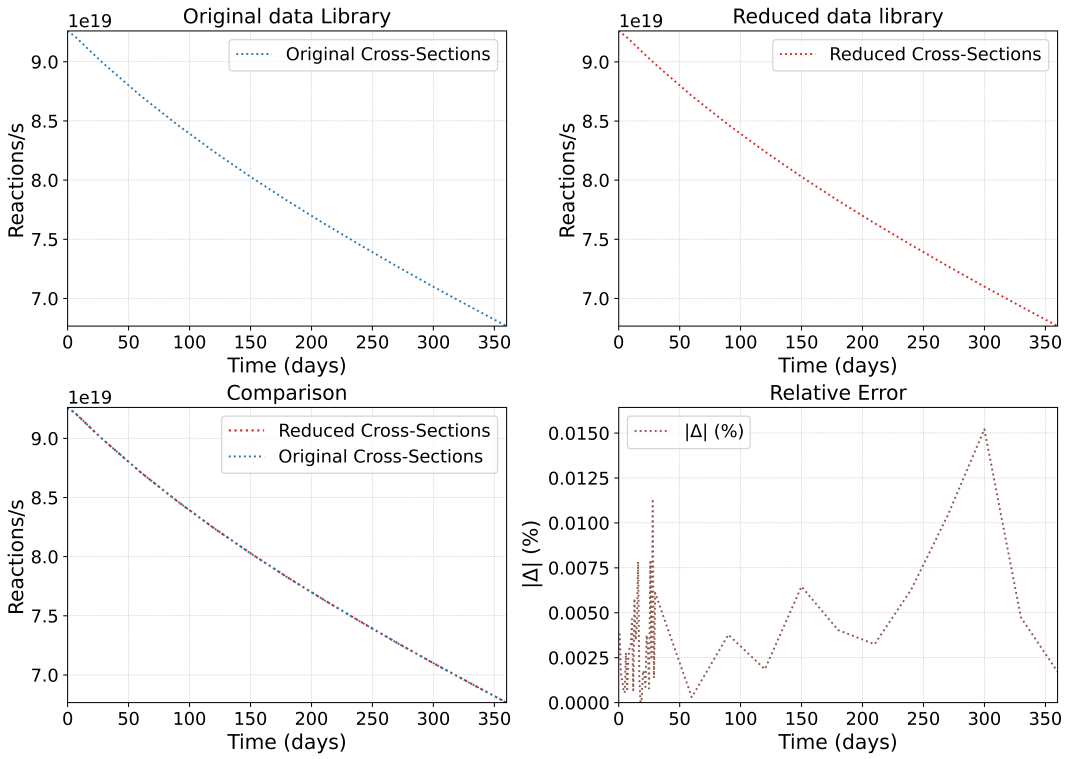


Figure 5. VVER-1000: U-235 (n, f) fission rate, original vs. reduced data libraries over 365 days.

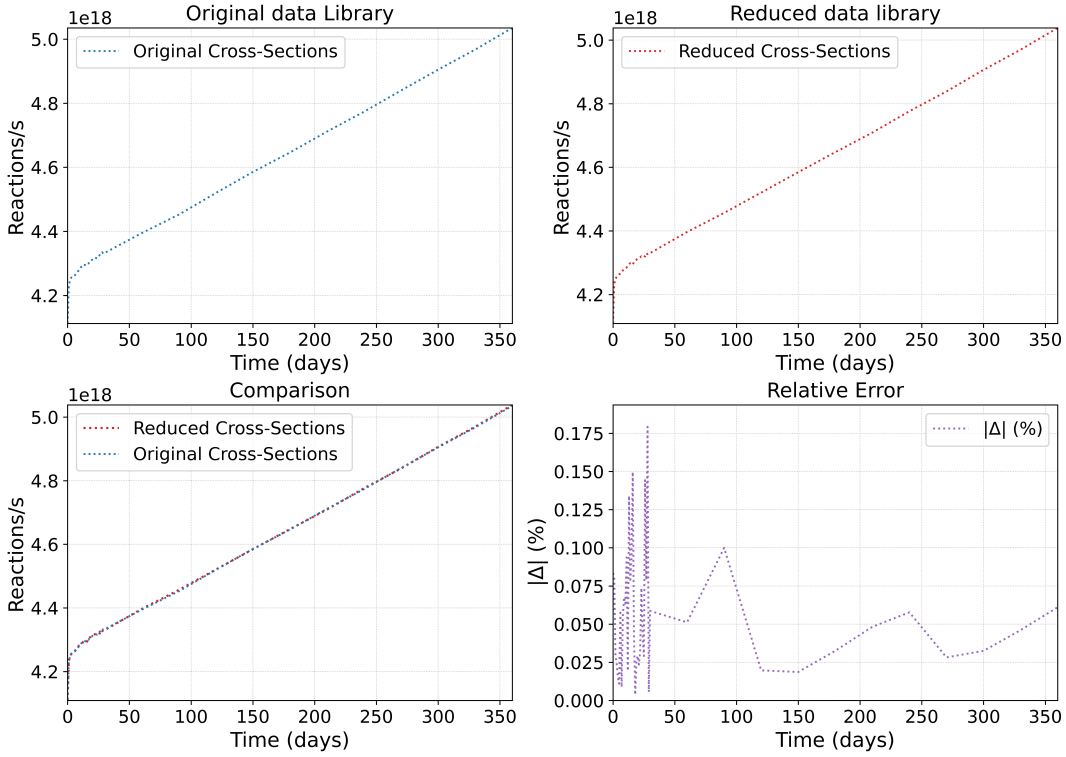


Figure 6. VVER-1000: U-238 (n, f) fission rate, original vs. reduced data libraries over 365 days.

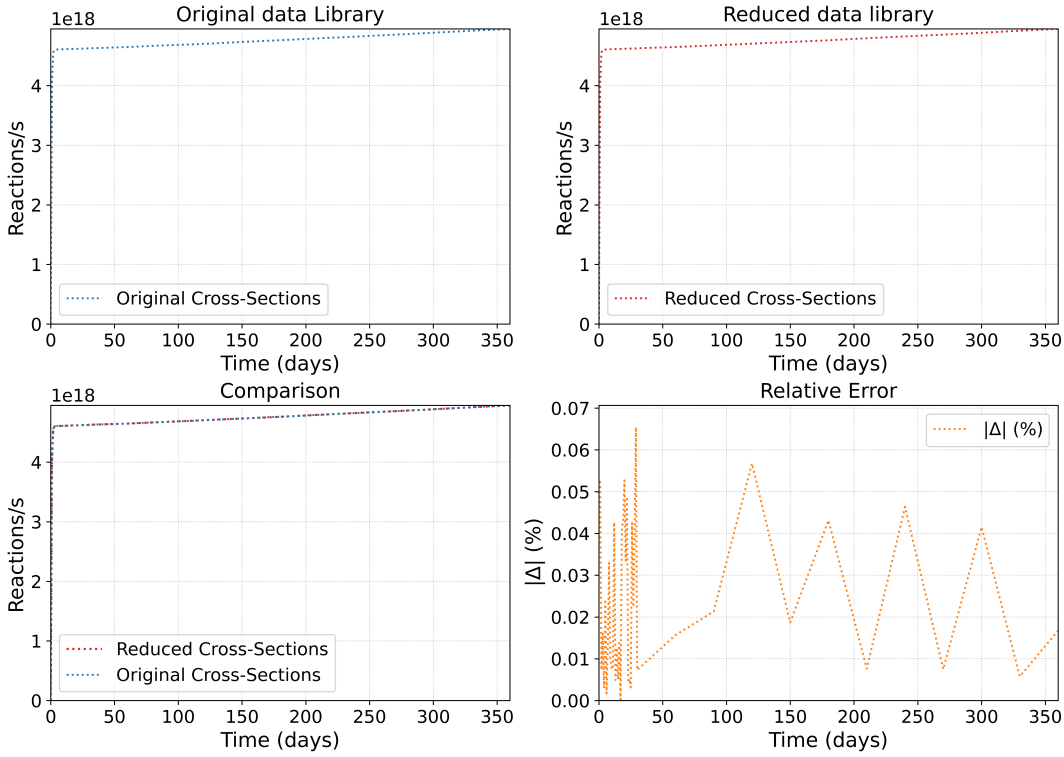


Figure 7. VVER-1000: Xe-135 (n, γ) capture rate, original vs. reduced data libraries over 365 days.

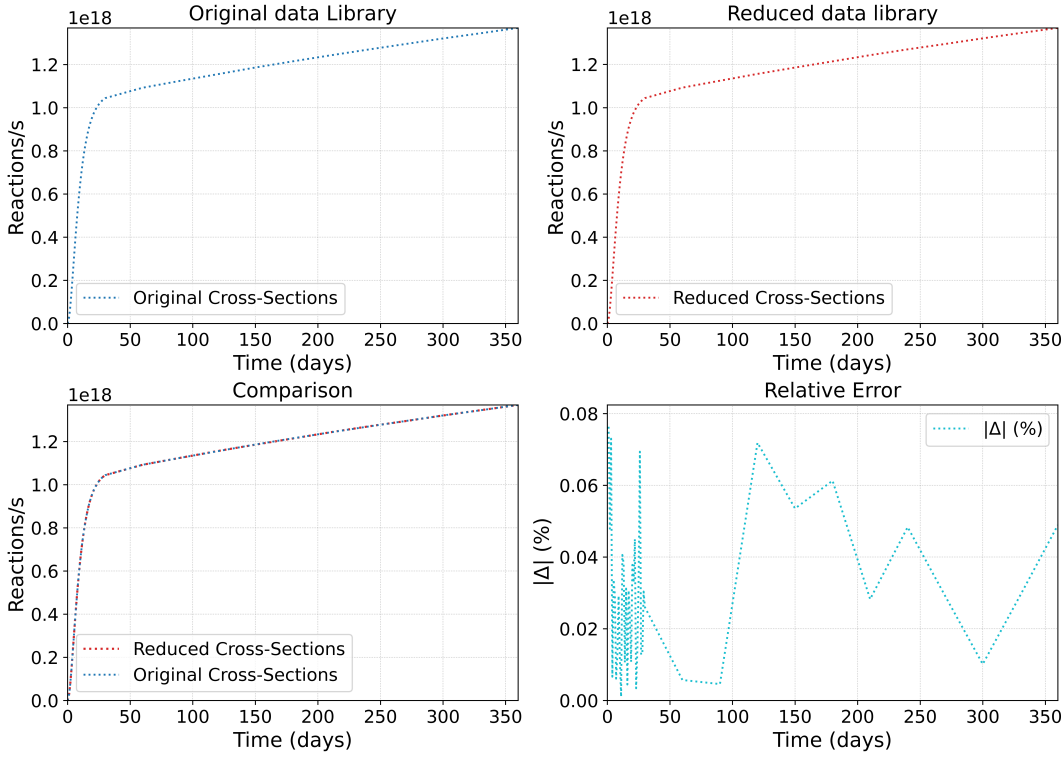


Figure 8. VVER-1000: Sm-149 (n, γ) capture rate, original vs. reduced data libraries over 365 days.

Isotopic inventories over depletion time

The inventories are examined for (i) the primary actinides U-235 and U-238, (ii) the dominant fission product poisons Xe-135 and Sm-149, and (iii) the plutonium vector Pu-238, Pu-239, Pu-240, Pu-241 and Pu-242. Figures 9, 10, 11, 12, 13, 14, 15, 16 and 17 display the original concentration vs. reduced time series. Table 7 provides quantitative summaries such as mean, 95th percentile, maximum (with time), and EOC relative differences over 365 days.

Isotope	Mean [%]	95th [%]	Max [%]	Day of max	EOC [%]
U-235	0.0005	0.0011	0.0039	356	0.0039
U-238	0.0001	0.0003	0.0004	338	0.0003
Xe-135	0.0183	0.0303	0.0353	30	0.0135
Sm-149	0.0240	0.0319	0.0650	153	0.0341
Pu-238	0.2487	0.3649	1.2868	11	0.0998
Pu-239	0.0497	0.0658	0.1285	17	0.0697
Pu-240	0.0443	0.0600	0.2123	1	0.1188
Pu-241	0.1331	0.1982	0.6040	14	0.0443
Pu-242	0.1301	0.2026	0.6202	13	0.0137

Table 7. VVER-1000: Relative absolute errors for isotopic concentrations over 365 days.

Primary actinides. For U-235 and U-238 the agreement is excellent. Over 365 days, the mean relative absolute differences are **0.0005%** for U-235 and **0.0001%** for U-238 (table 7). At the end of the cycle (EOC), the relative absolute differences are **0.0039%** for U-235 (day 356) and **0.0003%** for U-238 (day 338). These extremely small values indicate that the adaptive grid thinning preserved the integral balance between fission and capture on the principal actinides.

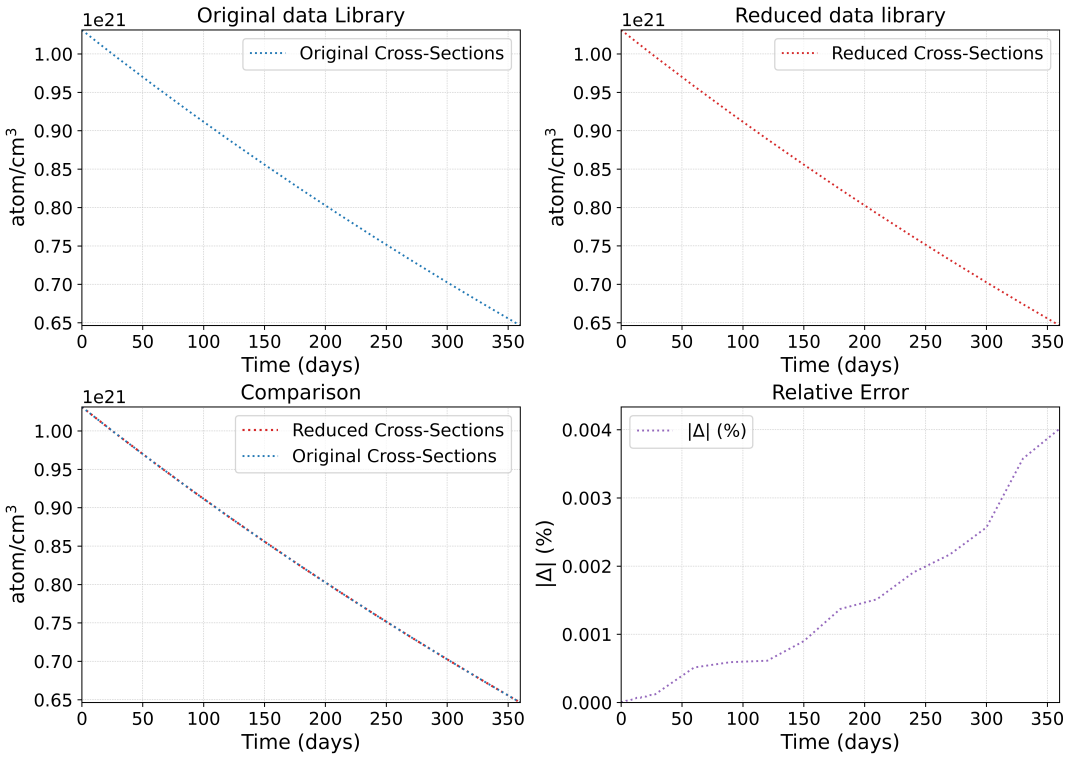


Figure 9. VVER-1000: U-235 concentration, original vs. reduced data libraries over 365 days.

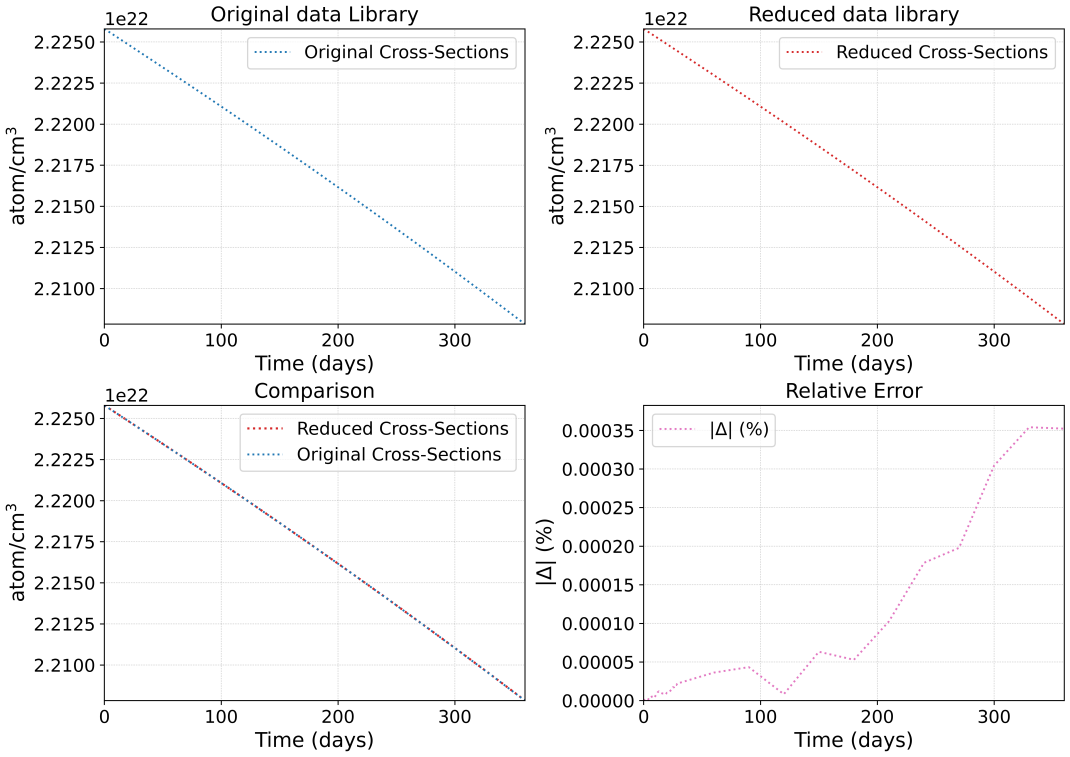


Figure 10. VVER-1000: U-238 concentration, original vs. reduced data libraries over 365 days.

The $\mathcal{O}(10^{-4}\text{–}10^{-3})\%$ EOC differences for U-235 and U-238 imply that the integrated balance between fission and capture on the principal actinides is essentially unchanged by thinning. In thermal PWRs, U-238(n,γ) in the epithermal range sets the feed to the Pu chain, while U-235(n,f) governs power normalization; preserving resonance-region detail ensures that both the capture-to-fission partition and the burn-up slope are maintained. The near-zero bias at EOC indicates that the spectral hardening with depletion has been tracked without distorting the epithermal absorption troughs that control U-238 self-shielding.

If U-238 capture had been materially perturbed, one would expect a correlated drift in the early formation of Pu-239 and in $k_{\text{eff}}(t)$ as plutonium worth emerges. Instead, the Pu-239 and Pu-241 EOC differences remain at **0.0697%** and **0.0443%** (table 7), and the k_{eff} EOC offset is only **11.12 pcm** (table 5). This coherence across inventories, reaction rates, and k_{eff} supports the conclusion that the reduction preserved the dominant actinide physics to within very tight tolerances.

Plutonium vector. Relative differences for the plutonium vector are mostly below one percent and are largest when inventories are very small, which inflates the relative metric. Over 365 days, the mean relative absolute differences are **0.2487%** for Pu-238, **0.0497%** for Pu-239, **0.0443%** for Pu-240, **0.1331%** for Pu-241, and **0.1301%** for Pu-242. The maxima occur early in the cycle when absolute concentrations are tiny, including a single excursion above one percent: **1.2868%** at day **11** for Pu-238; the other maxima are **0.1285%** at day **17** for Pu-239, **0.2123%** at day **1** for Pu-240, **0.6040%** at day **14** for Pu-241, and **0.6202%** at day **13** for Pu-242. By EOC, the relative absolute differences settle to **0.0998%** for Pu-238, **0.0697%** for Pu-239, **0.1188%** for Pu-240, **0.0443%** for Pu-241, and **0.0137%** for Pu-242.

This behavior is consistent with thermal PWR production pathways: U-238$(n,\gamma)\rightarrow$U-239 $(\beta^-)\rightarrow$ Np-239 $(\beta^-)\rightarrow$ Pu-239, followed by successive captures to higher Pu isotopes, while Pu-241 is fissile. Because inventories start near zero, relative metrics are most sensitive at beginning of life; as the plutonium vector builds in, differences stabilize to the small EOC values above. From a reactivity standpoint, the isotopes with the greatest worth, Pu-239 and Pu-241, show very small EOC differences **0.0697%** and **0.0443%**, consistent with the tight $|\Delta k|_{\text{pcm}}$ envelope reported in table 5. Also, the capture-dominated isotopes (Pu-240, Pu-242) also remain tightly controlled at EOC. This behavior indicates that the thresholds and resonance neighborhoods that control fertile capture and the first Pu in-growth are being sampled with sufficient fidelity. Together, these results suggest that the resonance-weighted thinning did not distort the fission/capture competition within the Pu family, a conclusion reinforced by the small, consistently negative k_{eff} shift observed across the cycle.

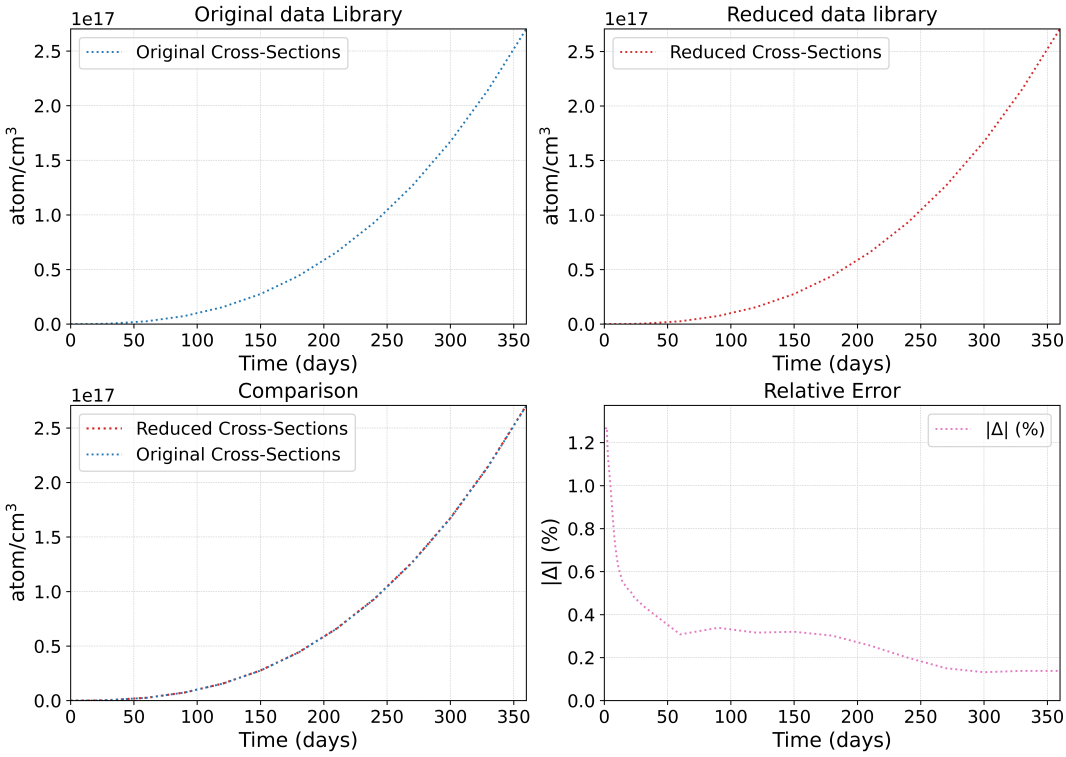


Figure 11. VVER-1000: Pu-238 concentration, original vs. reduced data libraries over 365 days.

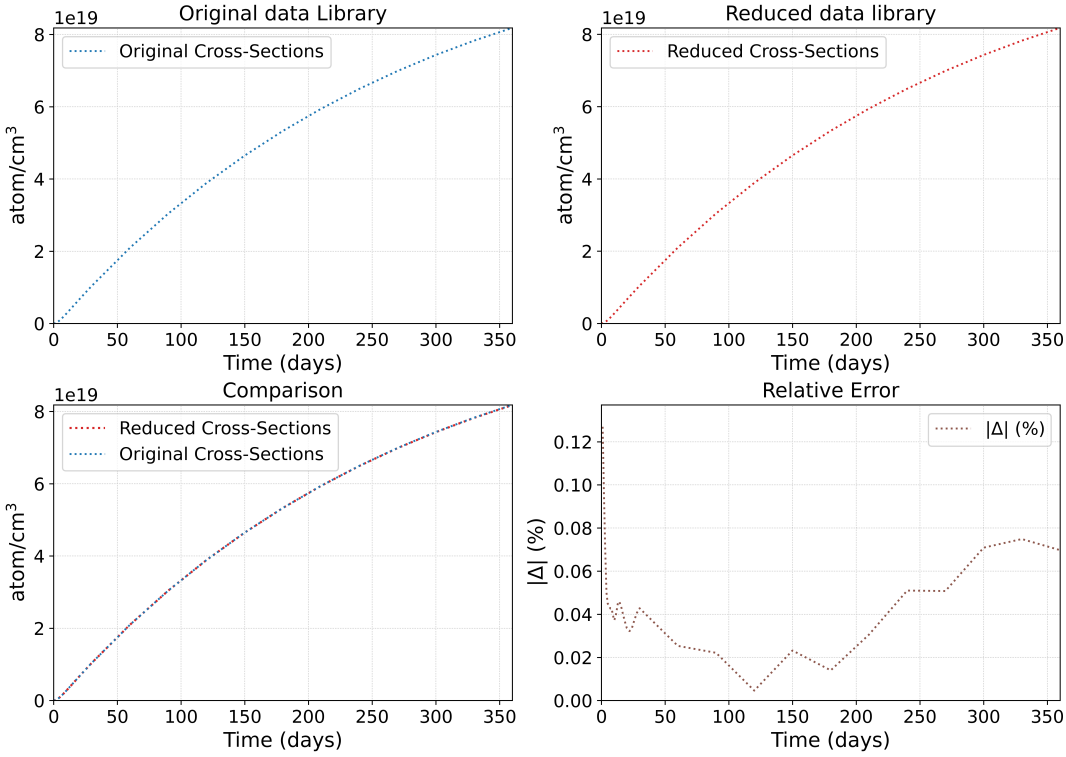


Figure 12. VVER-1000: Pu-239 concentration, original vs. reduced data libraries over 365 days.

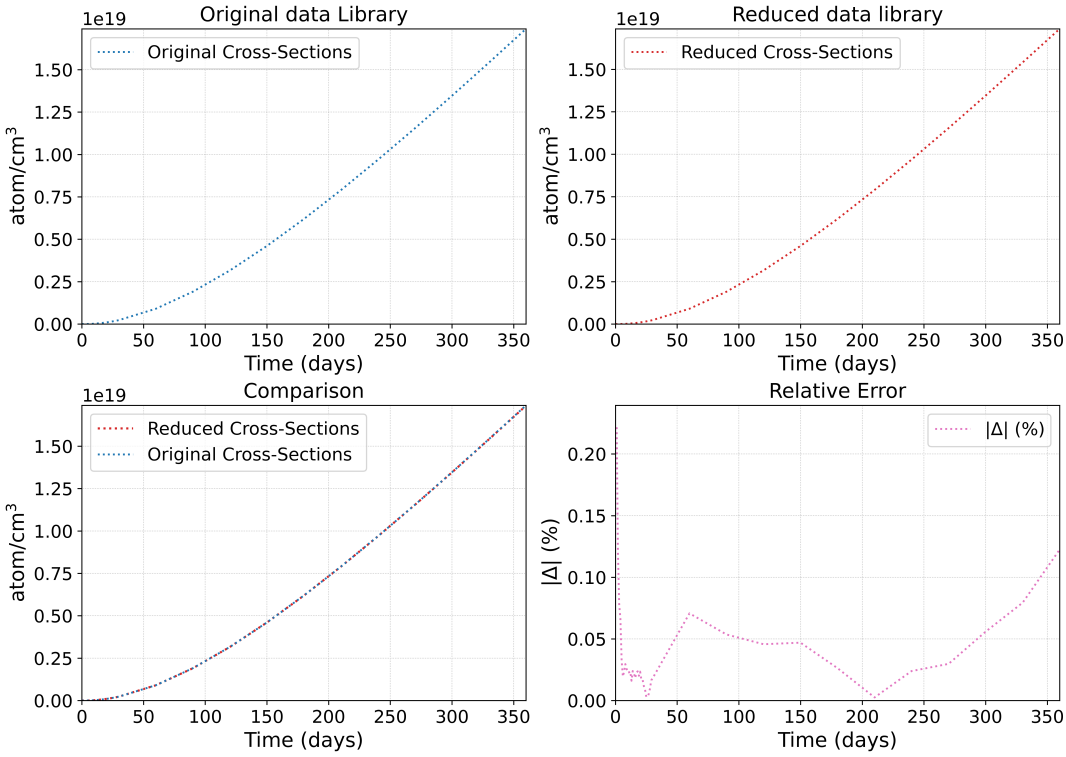


Figure 13. VVER-1000: Pu-240 concentration, original vs. reduced data libraries over 365 days.

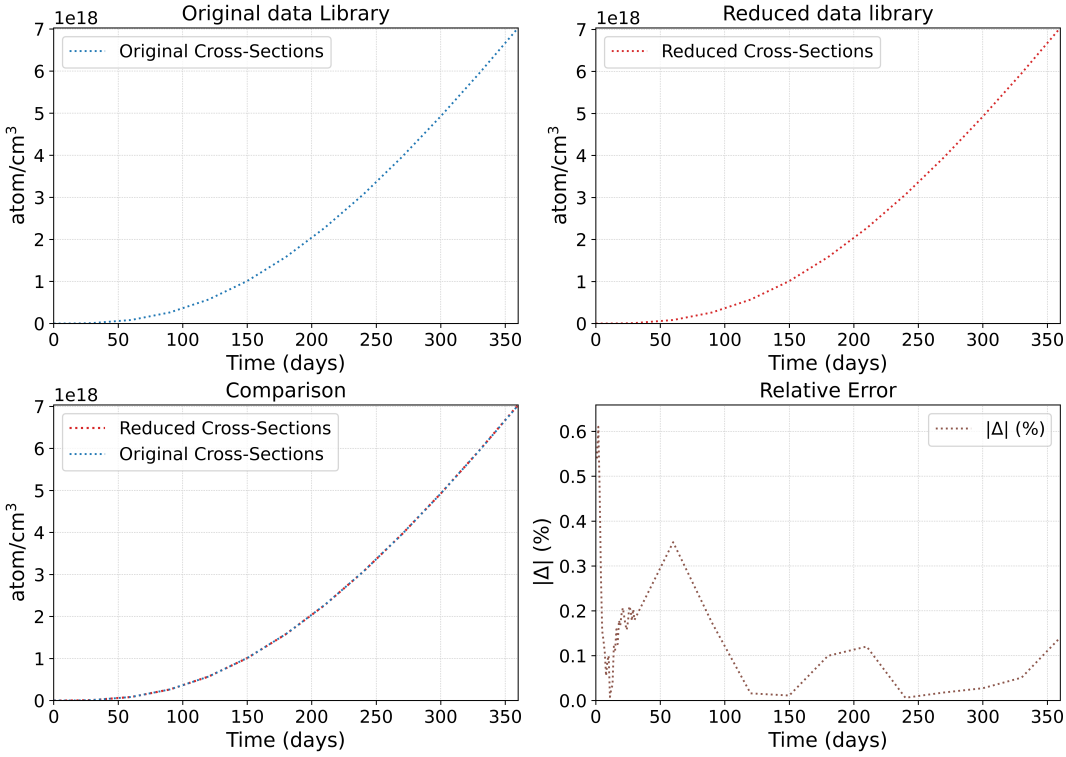


Figure 14. VVER-1000: Pu-241 concentration, original vs. reduced data libraries over 365 days.

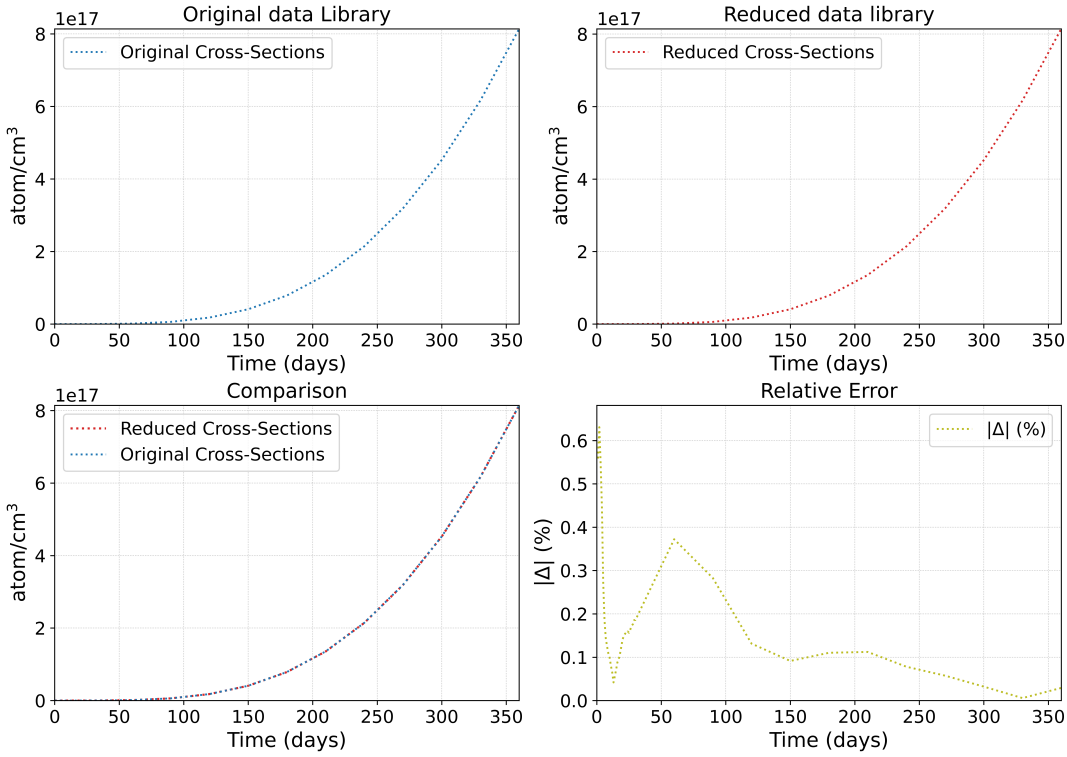


Figure 15. VVER-1000: Pu-242 concentration, original vs. reduced data libraries over 365 days.

Fission product poisons. The average relative differences are **0.0183%** for Xe-135 and **0.0240%** for Sm-149. The maxima are **0.0353%** at day **30** for Xe-135 and **0.0650%** at day **153** for Sm-149; the EOC values are **0.0135%** for Xe-135 and **0.0341%** for Sm-149. These small percentages are consistent with the high thermal absorption worth of xenon and samarium and their dynamic coupling to I-135→Xe-135 and Pm-149(β^-)→Sm-149.[13]

The early relative-difference peak for xenon is consistent with iodine–xenon transient kinetics at the start of irradiation; thereafter, xenon approaches an operating equilibrium while continuing to track slow spectral changes from fuel depletion. Sm-149 grows more gradually via Pm-149 (β^-) decay and saturates later, so a mid-cycle maximum is expected. The small EOC differences **0.0135%** Xe-135 and **0.0341%** Sm-149 indicate that thermal capture across the main resonance troughs has been preserved by the reduction, which is critical because these poisons set much of the short- and medium-term reactivity swing in PWRs.

Under the present assumptions (constant power, no control-rod motion, fixed chemistry), xenon and samarium remain within tight error bands. In more dynamic scenarios—e.g. daily boron swing, rod insertion, or temperature transients—the thermal flux shape may change more abruptly. Because the reduction explicitly preserves thresholds and emphasizes resonance regions, similar behavior is expected, but transient cases should be checked separately to confirm that xenon and samarium feedbacks remain within decision-relevant bounds.

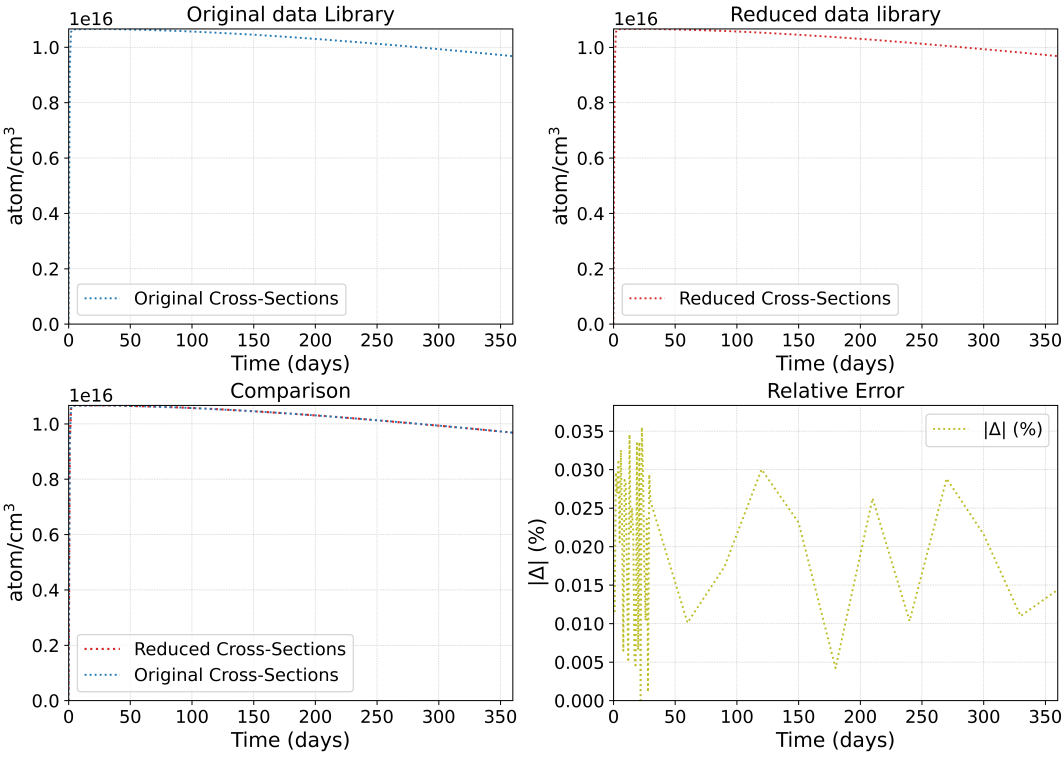


Figure 16. VVER-1000: Xe-135 concentration, original vs. reduced data libraries over 365 days.

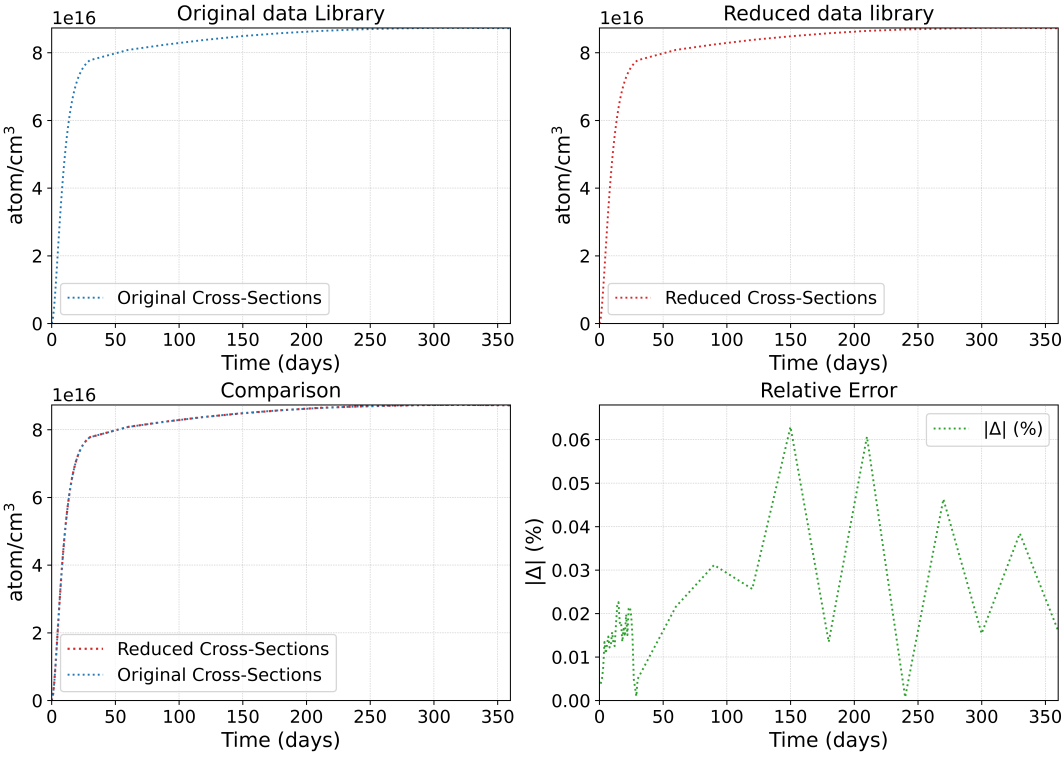


Figure 17. VVER-1000: Sm-149 concentration, original vs. reduced data libraries over 365 days.

Core Physics behavior

Over one year of depletion in a full-core of VVER-1000 with all other OpenMC's acceleration features disabled (WMP=false, default equal-lethargy energy mapping), by using the ML-thinned pointwise library, the results show;

- k_{eff} **remains within 96.79 pcm of the baseline at all times.** Using the reduced library, the absolute difference $|\Delta k|_{\text{pcm}}$ over the cycle has *mean* **29.21 pcm**, *95th-percentile* **55.36 pcm**, *maximum* **96.79 pcm** at day **12**, and *EOC* **11.12 pcm** (table 5). The small, persistent down-shift is consistent with a slight re-weighting of reaction competition under grid thinning while thresholds and resonance regions are preserved.

- **Key reaction rates are preserved to $\leq 0.18\%$ with physically consistent timing.** Over the cycle, the mean absolute relative differences are **0.0036%** (U-235 (n,f)), **0.0556%** (U-238 (n,f)), **0.0232%** (Xe-135 (n,γ)), and **0.0299%** (Sm-149 (n,γ)). Maxima are all below **0.1790%** and occur at day **300**, **28**, **29**, and **210**, respectively (table 6). The timing matches reactor-physics expectations (early xenon transient; BOL fast-fission prominence of U-238; mid-cycle Pm-149(β^-)$\rightarrow$Sm-149 feeding; late U-235 fission).

- **Actinide and poison inventories track to $\ll 1\%$ throughout.** At EOC the relative absolute differences are **0.0039%** (U-235), **0.0003%** (U-238), **0.0135%** (Xe-135), and **0.0341%** (Sm-149), with cycle-averaged means of **0.0005%**, **0.0001%**, **0.0183%**, and **0.0240%**, respectively (table 7).

- **Plutonium-vector differences are sub-0.12% at EOC.** Early-time maxima—e.g., **1.2868%** at day **11** for Pu-238—arise when inventories are tiny; by EOC the relative absolute differences reduce to **0.0998%** (Pu-238), **0.0697%** (Pu-239), **0.1188%** (Pu-240), **0.0443%** (Pu-241), and **0.0137%** (Pu-242) (table 7). The isotopes with the strongest reactivity leverage (Pu-239, Pu-241) therefore remain tightly controlled, consistent with the small $|\Delta k|_{\text{pcm}}$ envelope.

The comparisons above are *deterministic* differences between two libraries under identical Monte Carlo settings; per-step stochastic uncertainties were not recorded. Relative error spikes at early times (small denominators) or near resonance troughs are expected; we therefore interpret physics significance using EOC values together with time-averaged statistics.
Overall, the machine-learned reduction of pointwise nuclear data for the targeted 23 nuclides yields meaningful wall-time and memory benefits (Table 4) while keeping full-core criticality, reaction rates, and inventories within narrow, quantified bounds (*e.g.*, mean $|\Delta k|_{\text{pcm}}$ **29.21**, maxima $<$ **100** pcm; reaction-rate maxima $\leq$ **0.1790%**; EOC inventory differences $\leq$ **0.1188%**).

Conclusions

This study examines the impact of continuous-energy nuclear data reduction by combining machine learning driven adaptive resampling with direct HDF5 modification on a one-year Monte Carlo full-core criticality and burn-up analysis using OpenMC. The reduction targeted 23 nuclides effective on the pressurized water reactors (PWR) reactivity and burn-up, retained roughly 10%–50% of energy grid and cross-sections per nuclide while explicitly preserving the thresholds and resonance structure. To isolate the effect of the developed reduction method, the windowed multipole method has been disabled.
Two industrial scale full-core EPR and VVER-1000 models are used to benchmark the reduced nuclear data library against the original one in terms of computational performance and fidelity. Using the reduced nuclear data library, wall-clock time decreased by **17.81%** in EPR and **42.5%** in VVER-1000. MaxRSS decreased by **4.4%** in EPR and increased by **5.0%** in VVER-1000, respectively. Relative errors remained bounded. For VVER-1000 over a one-year burn-up, the $k_{\text{eff}}(t)$ difference (reported in pcm) had a **mean of 29.21 pcm**, a **95th-percentile of 55.36 pcm**, a **maximum of 96.79 pcm at day 12**, and an **EOC value of 11.12 pcm**.
Relative differences in reaction rates remained sub-percent throughout. Over the year, the temporal means were **0.0036%** for U-235 (n,f), **0.0556%** for U-238 (n,f), **0.0232%** for Xe-135 (n,γ), and **0.0299%** for Sm-149 (n,γ). The maxima occurred at reactor-physically plausible times: **0.0150%** at day **300** for U-235 (n,f), **0.1790%** at day **28** for U-238 (n,f), **0.0656%** at day **29** for Xe-135 (n,γ), and **0.0763%** at day **210** for Sm-149 (n,γ). The corresponding EOC relative differences were **0.0017%** (U-235 (n,f)), **0.0605%** (U-238 (n,f)), **0.0128%** (Xe-135 (n,γ)), and **0.0300%** (Sm-149 (n,γ)).
Isotopic inventories at EOC were preserved at very small relative differences: **0.0039%** for U-235 and **0.0003%** for U-238, **0.0135%** for Xe-135, and **0.0341%** for Sm-149. Across the plutonium vector, the EOC relative differences were **0.0998%** (Pu-238), **0.0697%** (Pu-239), **0.1188%** (Pu-240), **0.0443%** (Pu-241), and **0.0137%** (Pu-242). Early-time spikes in the Pu isotopes occur when absolute inventories are very small and thus inflate relative metrics; by EOC the differences settle to the values reported here.
Taken together, the resonance and energy region aware reduction developed method following by reduction in energy grid and cross-section of a targeted nuclide set can yield meaningful performance gains while keeping full-core criticality, reaction rates, and inventories within narrow, quantifiable error envelopes. The conclusions pertain to continuous-energy cross-sections

transport with ENDF/B-VII.1, constant-power depletion, and no control-rod motion; per-step Monte Carlo uncertainties were not recorded, so differences were evaluated deterministically. These results indicate that the developed reduction method accelerates full-core analysis and reduces MaxRSS while maintaining fidelity in neutronics studies.
In forthcoming studies, several extensions are natural. Methodological work could follow to enable WMP and reassess the accuracy–performance trade space, compare directly with NJOY/RECONR at matched tolerances, propagate evaluated covariances to quantify uncertainty under thinning, and study interactions with alternative grid search or unionized grid strategies. Scope of studies could extend to repeat with ENDF/B-VIII.0[22], JEFF-3.3[23], and JENDL[24] libraries, test other fission reactor types (SFR, LFR, HTGR, MSR, Heavy Water CANDU, BWR) and transient scenarios (Xenon dynamics, control-rod insertion, coupled TH feedback), as well as fusion reactors design including evaluating non-reactor applications such as detector design, shielding, fusion materials at 14 MeV. In term of portability studies, it could be extended to apply the workflow reduction of nuclear data method to other Monte Carlo codes such as Serpent[25], MCNP[26], SCALE[27], Geant4[28], PHITS[29] via their nuclear data interfaces. These studies will help map the operational envelope in which machine learned nuclear data reduction delivers robust acceleration with controlled, decision-relevant biases.

Acknowledgements

The authors of this paper gratefully acknowledge the valuable scientific guidance provided by Dr. Paul Romano and Dr. Patrick C. Shriwise from the OpenMC team, whose insights greatly contributed to the success of this work.

Author contributions statement

Alexander Hashemi was responsible for the design and implementation of the code and algorithms, carried out the simulations, and prepared all relevant materials. Rafael Macian-Juan and Martin Ohlerich provided project supervision. All authors reviewed and approved the final manuscript.

Corresponding author

Correspondence to Alexander Hashemi.

Data availability

The data used and/or analysed during the current study, including simulation outputs and model predictions are available from the corresponding author on reasonable request.

Additional information

The data used and/or analysed during the current study, including simulation outputs and model predictions are available from the corresponding author on reasonable request.

Funding statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interests

All authors confirm that there are no competing interests in relation to this work.

References

1. Romano, P. K. & et al. Openmc: A state-of-the-art monte carlo code for research and development. *Annals Nucl. Energy* **82**, 90–97, DOI: https://doi.org/10.1016/j.anucene.2014.07.048 (2015).
2. Folk, M., Heber, G., Koziol, Q., Pourmal, E. & Robinson, D. An overview of the hdf5 technology suite and its applications. In *Proceedings of the EDBT/ICDT 2011 workshop on array databases*, 36–47, DOI: https://doi.org/10.1145/1966895.1966900 (2011).
3. Macfarlane, R., Muir, D. W., Boicourt, R. M., Kahler, A. C. & Conlin, J. L. The njoy nuclear data processing system, version 2016. Tech. Rep., Los Alamos National Laboratory (LANL), Los Alamos, NM (United States) (2017). DOI: https://doi.org/10.2172/1338791.

4. Chadwick, M. *et al.* Endf/b-vii.1 nuclear data for science and technology: Cross sections, covariances, fission product yields and decay data. *Nucl. Data Sheets* **112**, 2887–2996, DOI: https://doi.org/10.1016/j.nds.2011.11.002 (2011). Special Issue on ENDF/B-VII.1 Library.
5. Brown, F. B. New hash-based energy lookup algorithm for monte carlo codes. Tech. Rep. LA-UR-14-24530, Los Alamos National Laboratory (2014).
6. Openmc documentation: Cross section representations. https://docs.openmc.org/en/stable/methods/cross_sections.html. Accessed 2025-09-13.
7. Josey, C., Ducru, P., Forget, B. & Smith, K. Windowed multipole for cross section doppler broadening. *J. Comput. Phys.* **307**, 715–727, DOI: https://doi.org/10.1016/j.jcp.2015.08.013 (2016).
8. Hashemi, A., Macián-Juan, R. & Ohlerich, M. Development of a novel machine learning-based adaptive resampling algorithm for nuclear data processing. *Sci. Reports* **15**, 32573, DOI: https://doi.org/10.1038/s41598-025-18674-8 (2025).
9. Openmc documentation: Neutron physics. https://docs.openmc.org/en/stable/methods/neutron_physics.html. Accessed 2025-09-13.
10. Openmc documentation: Depletion chains representations. https://openmc.org/depletion-chains/ , https://docs.openmc.org/en/stable/pythonapi/generated/openmc.deplete.Chain.html. Accessed 2025-09-13.
11. Openmc documentation: Depletion. https://docs.openmc.org/en/stable/methods/depletion.html. Accessed 2025-09-13.
12. Openmc documentation: Depletion methods. https://docs.openmc.org/en/stable/pythonapi/deplete.html. Accessed 2025-09-13.
13. Doe fundamentals handbook: Nuclear physics and reactor theory, volume 2 (doe-hdbk-1019/2-93). https://www.standards.doe.gov/standards-documents/1000/1019-bhdbk-1993-v2 (1993). Accessed 2025-09-13.
14. Openmc documentation: Windowed multipole representation. https://docs.openmc.org/en/stable/pythonapi/generated/openmc.data.WindowedMultipole.html. Accessed 2025-09-13.
15. Framatome ANP, Inc. Epr design description. Technical Report ML052280170, U.S. Nuclear Regulatory Commission, Rockville, MD, USA (2005). Vendor design description submitted to NRC.
16. Leskin, S. T., Shelegov, A. S. & Slobodchuk, V. I. *Fizicheskie osobennosti i konstruktsiya reaktora VVER-1000 [Physical Features and Design of the VVER-1000 Reactor]* (National Research Nuclear University "MEPhI", Moscow, 2011). In Russian.
17. Harris, C. R. & et al. Array programming with *numpy*. *Nature* **585**, 357–362, DOI: https://doi.org/10.1038/s41586-020-2649-2 (2020).
18. Virtanen, P. & et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272, DOI: https://doi.org/10.1038/s41592-019-0686-2 (2020).
19. Collette, A. *Python and HDF5: unlocking scientific data* (" O'Reilly Media, Inc.", 2013).
20. Pedregosa, F. & et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
21. Hunter, J. D. Matplotlib: A 2d graphics environment. *Comput. Sci. & Eng.* **9**, 90–95, DOI: https://doi.org/10.1109/MCSE.2007.55 (2007).
22. Brown, D. *et al.* Endf/b-viii.0: The 8th major release of the nuclear reaction data library with cielo-project cross sections, new standards and thermal scattering data. *Nucl. Data Sheets* **148**, 1–142, DOI: https://doi.org/10.1016/j.nds.2018.02.001 (2018). Special Issue on Nuclear Reaction Data.
23. Plompen, A. J. *et al.* The joint evaluated fission and fusion nuclear data library, jeff-3.3. *The Eur. Phys. J. A* **56**, 1–108 (2020).
24. Iwamoto, O. *et al.* Japanese evaluated nuclear data library version 5: Jendl-5. *J. Nucl. Sci. Technol.* **60**, 1–60, DOI: https://doi.org/10.1080/00223131.2022.2141903 (2023).
25. Leppänen, J., Pusa, M., Viitanen, T., Valtavirta, V. & Kaltiaisenaho, T. The serpent monte carlo code: Status, development and applications in 2013. *Annals Nucl. Energy* **82**, 142–150, DOI: https://doi.org/10.1016/j.anucene.2014.08.024 (2015).
26. Kulesza, J. A. e. a. Mcnp® code version 6.3.0 theory & user manual. Tech. Rep., Los Alamos National Laboratory (LANL), Los Alamos, NM (United States) (2022). DOI: https://doi.org/10.2172/1889957.
27. Wieselquist, W. A. & Lefebvre, R. A. Scale 6.3.2 user manual. Tech. Rep., Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States) (2024). DOI: https://doi.org/10.2172/2361197.

28. Agostinelli, S. *et al.* Geant4—a simulation toolkit. *Nucl. Instruments Methods Phys. Res. Sect. A: Accel. Spectrometers, Detect. Assoc. Equip.* **506**, 250–303, DOI: https://doi.org/10.1016/S0168-9002(03)01368-8 (2003).

29. Sato, T. *et al.* Features of particle and heavy ion transport code system (phits) version 3.02. *J. Nucl. Sci. Technol.* **55**, 684–690, DOI: https://doi.org/10.1080/00223131.2017.1419890 (2018).