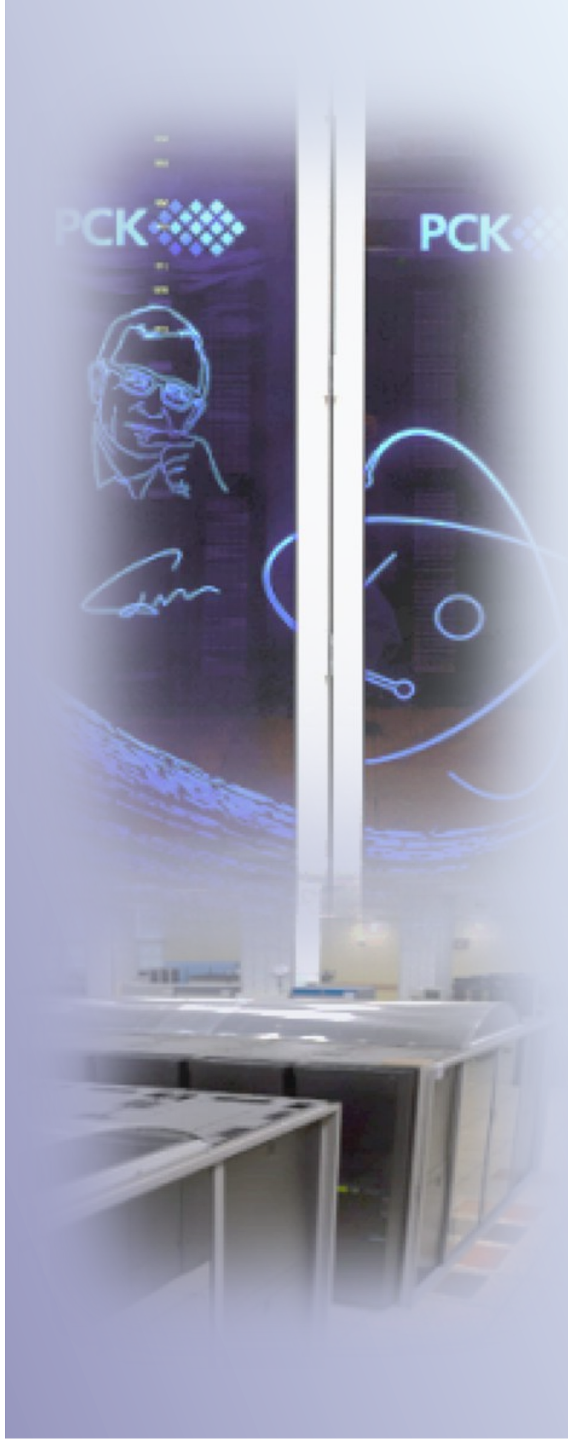
Random Forest Regression and Shapley Additive Explanation for Effective Dose Rate Estimation in High-Energy Neutron Fields Based on Bonner Spectrometer Measurements

Chizhov K.A.[1,2]

[1]Joint Institute for Nuclear Research,
Meshcheryakov Laboratory of Information Technologies, Dubna, Russia.

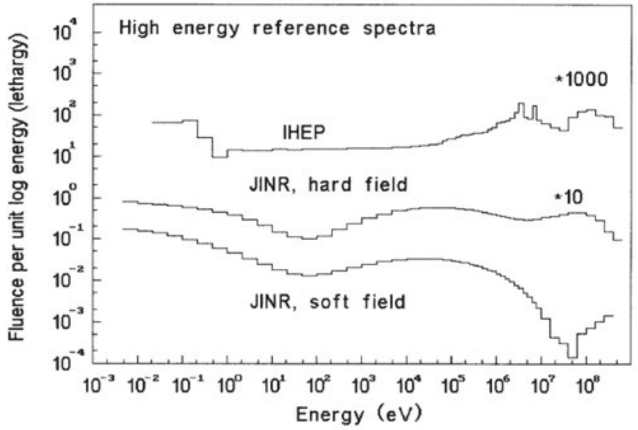
[2]University “Dubna”, Dubna, Russia.

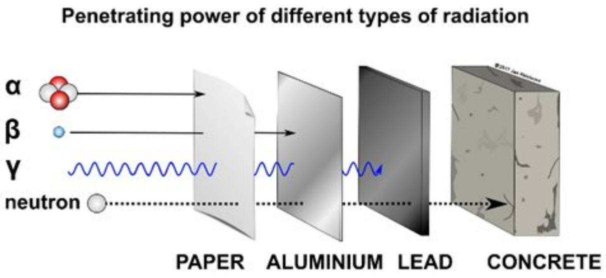


The International Conference dedicated to mathematics in artificial intelligence,
Sirius International Mathematics Center, March, 24-28, 2025

Introduction

- Individual dosimetric occupational monitoring is based on the results of dosimetric monitoring of workplaces, which consists of determining the energy distribution of **neutron flux density** when monitoring continuous radiation fields, as well as the time a worker spends in these conditions.
- Radiation fields behind the protective shields of the JINR nuclear physics facilities (particle accelerator, nuclear reactors) are formed mainly by **neutrons** of a wide energy spectrum.
- Radiation monitoring at accelerators cannot be carried out using standard dosimeters and neutron radiometers alone, since their operating range is limited by a maximum neutron energy of about 10-20 MeV.
- **Bonner Multisphere Spectrometer** is a common used tool for radiation monitoring at accelerators.

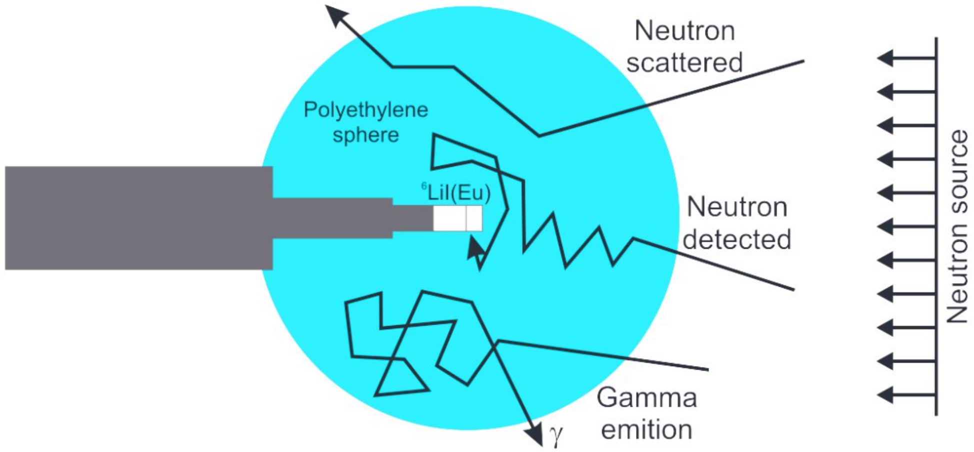




Bonner Multi-sphere Spectrometer

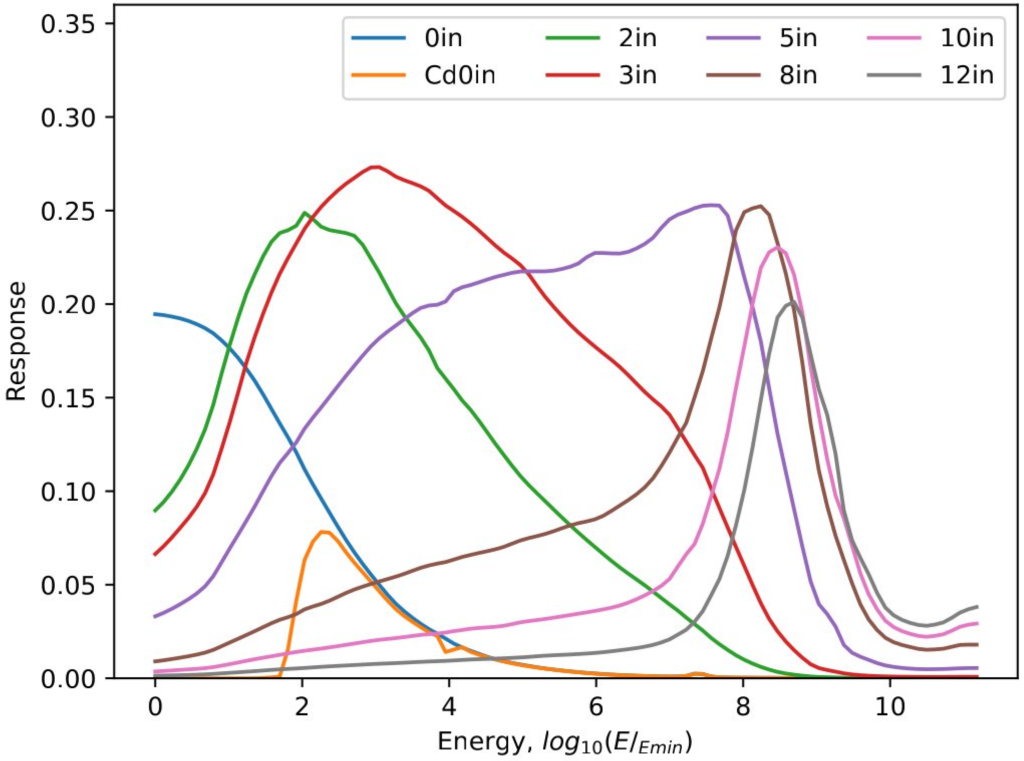
- A multi-sphere Bonner spectrometer is used to measure the neutron flux density.
- The measurement method is based on the moderation of fast neutrons in polyethylene moderator spheres of various diameters.
- Various detectors are used to detect thermal neutrons, e.g., inorganic scintillators such as ^{6}LiI.

https://atomtex.com/

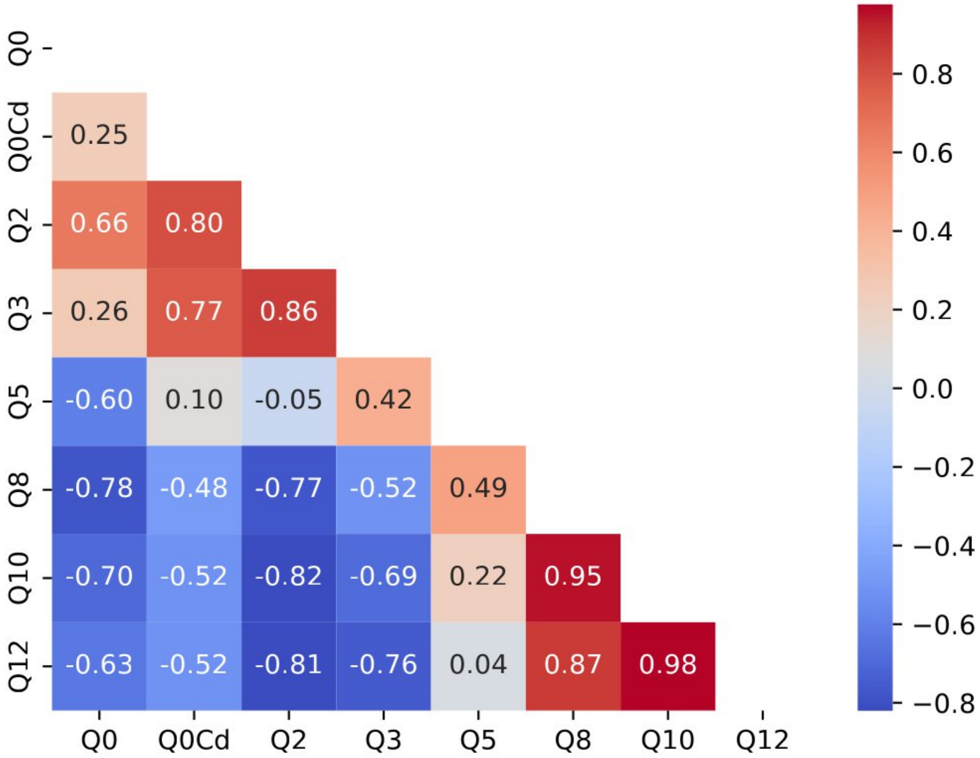


http://dx.doi.org/10.5772/51274

Response functions of moderator spheres



Martinkovic, J., and G. N. Timoshenko. *Calculation of multisphere neutron spectrometer response functions in energy range up to 20 MeV*. No. JINR-R--16-2005-105. Division of Radiation and Radiobiological Research, 2005.



Unfolding the neutron spectra

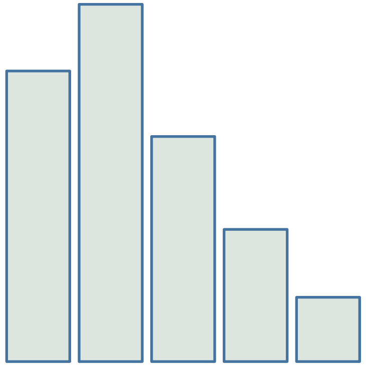
System of Fredholm integral equations of the 1st kind:

$$\begin{cases} \int_{E_{\min}}^{E_{\max}} K_1(E)\,\varphi(E)\,dE = Q_1, \\ \qquad\qquad \vdots \\ \int_{E_{\min}}^{E_{\max}} K_M(E)\,\varphi(E)\,dE = Q_M, \end{cases}$$

where:

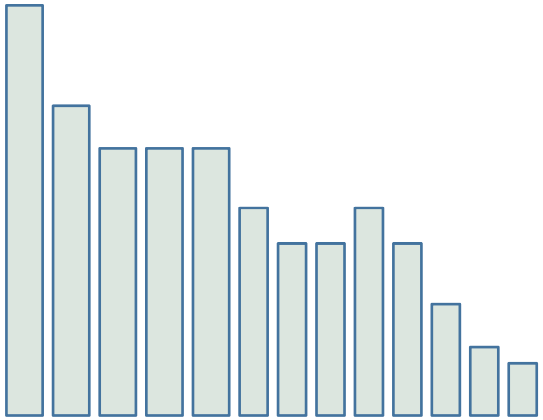
- Q_j is the Bonner spectrometer reading for the j-th sphere,
- $\varphi(E)$ - neutron spectrum,
- $K_j(E)$ is the kernel of the j-th equation, which is a response function of the detector to neutrons of various energies,
- M is the number of spheres used to measure the spectrum.
- The integration limits $E_{\min}$ and $E_{\max}$ are specified by the domain of definition of the neutron spectrum E and the set of detectors used for measurements.

Direct and inverse problems of spectrum unfolding

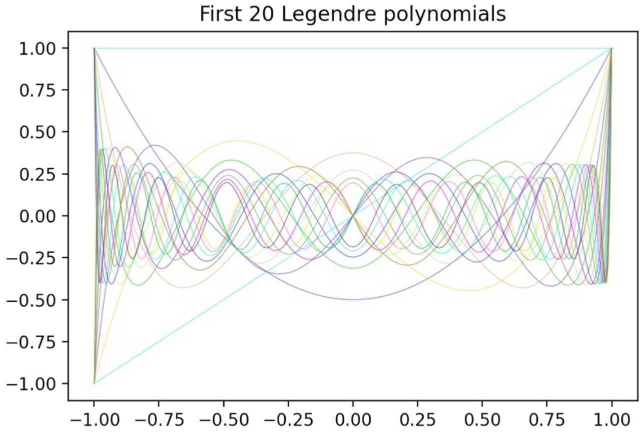


$Q_1 \ldots Q_M$

$M < S$
$M < N$



$E_1 \ldots E_S$



$C_1 \ldots C_N$

$$Q_j \approx \sum_{i=1}^{S} K_{ji} \Phi(\mathrm{E}_i) \Delta E, \quad j = 1,\ldots,M$$

$$\Phi = \sum_{i=1}^{N} C_i \, P_{i-1}$$

- Direct task: to obtain “effective” readings of the Bonner spectrometer based on the given spectra.
- Inverse task: to obtain the initial spectra based on the measurement results.

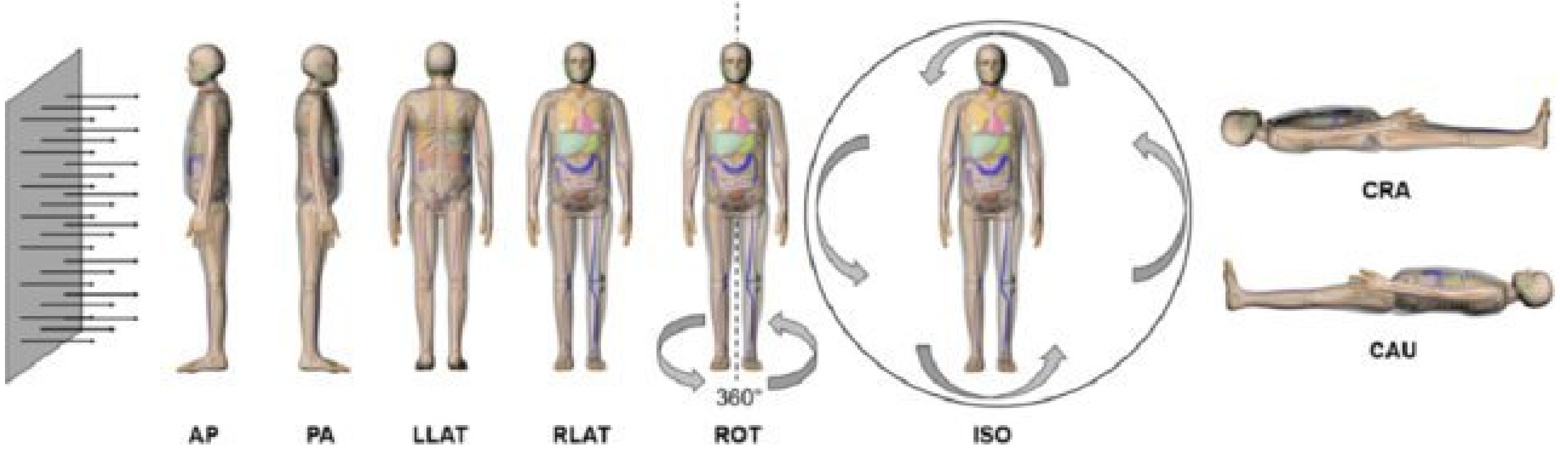
Dose estimation for spectrum

The Bonner multi-sphere spectrometer is used to measure neutron spectra in stationary fields to assess irradiation of personnel.

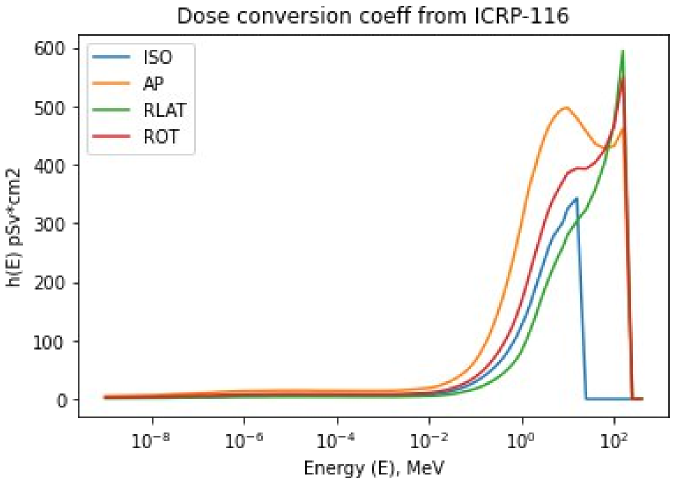
$$\dot{H} = \int_{Emin}^{Emax} \varphi(E) \cdot h(E) dE$$

$\dot{H}$ – dose rate ($\dot{E}_{\text{eff\_AP}}$,$\dot{E}_{\text{eff\_ISO}}$, $\dot{H}$*(10), $\dot{H}_p(10,0^{\circ})$)

h(w) – corresponding conversion factor [pSv·cm2] for monoenergetic particles in different irradiation geometries, ICRP Publication 116*



https://dceg.cancer.gov/tools/radiation-dosimetry-tools/phantoms-library

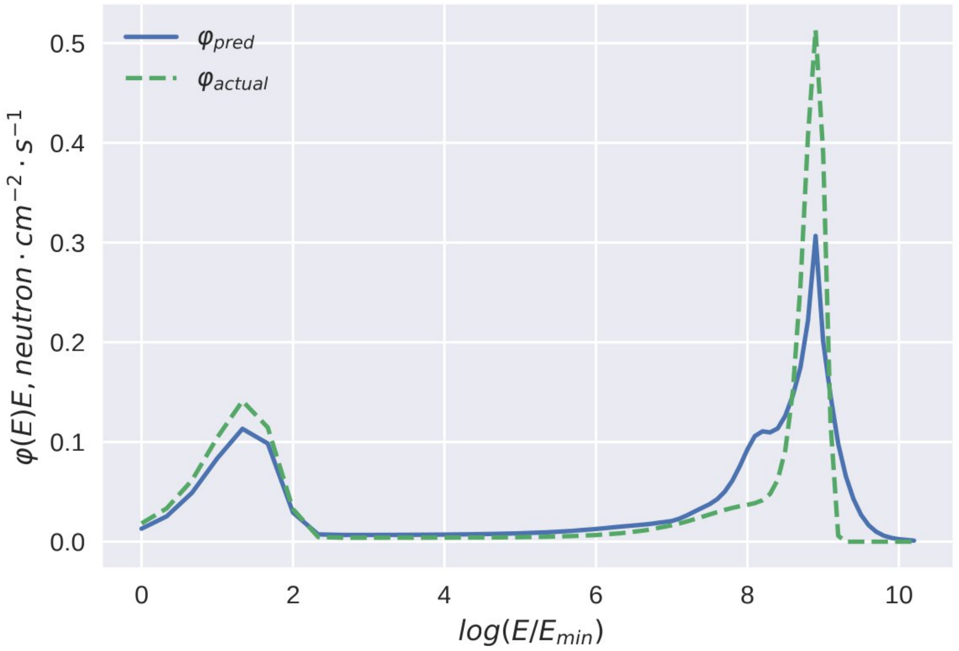


*Petoussi-Henss, Nina, et al. "ICRP Publication 116—the first ICRP/ICRU application of the male and female adult reference computational phantoms." *Physics in Medicine & Biology* 59.18 (2014): 5209.

Previously developed methods

1. Regularization (Tikhonov,...)
2. Statistical regularization
3. Maximum entropy principle
4. Maximum likehood
5. Genetic
6. Iterative
7. Parametric
8. Bayesian parameter estimation
9. Machine learning + neural networks
10. Combinations of methods

11....



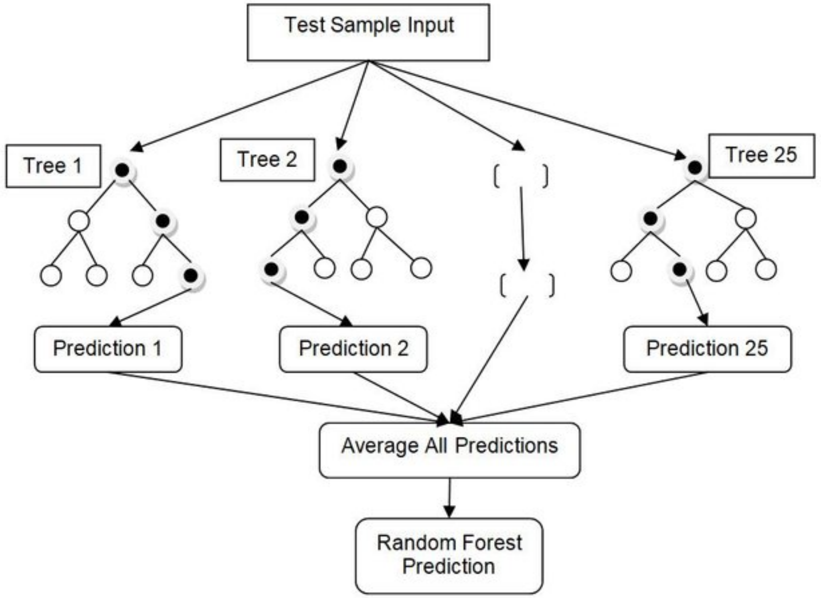
- Gómez-Ros, J. M., et al. "Results of the EURADOS international comparison exercise on neutron spectra unfolding in Bonner spheres spectrometry." *Radiation Measurements* 153 (2022): 106755.
- Barros, S., et al. "Comparison of unfolding codes for neutron spectrometry with Bonner spheres." *Radiation protection dosimetry* 161.1-4 (2014): 46-52.
- Chizhov, K., L. Beskrovnaya, and A. Chizhov. "Neutron Spectra Unfolding from Bonner Spectrometer Readings by the Regularization Method Using the Legendre Polynomials." *Physics of Particles and Nuclei* 55.3 (2024): 532-534.

Methodology

ML method & implementation

The random forest machine learning algorithm is a model consisting of an ensemble of decision trees. In a regression problem, the method is used to predict numerical values, the responses of each tree are averaged, while when constructing trees, a random sample of elements from the data set occurs, and when dividing nodes, random sets of parameters are selected.

- Implementation in *Python* + *Sklearn* + *optuna* + *NumPy* + *pandas* + *matplotlib*;
- Calculation on the multifunctional information and computing complex of the JINR Information Technology Laboratory, http://hlit.jinr.ru/.

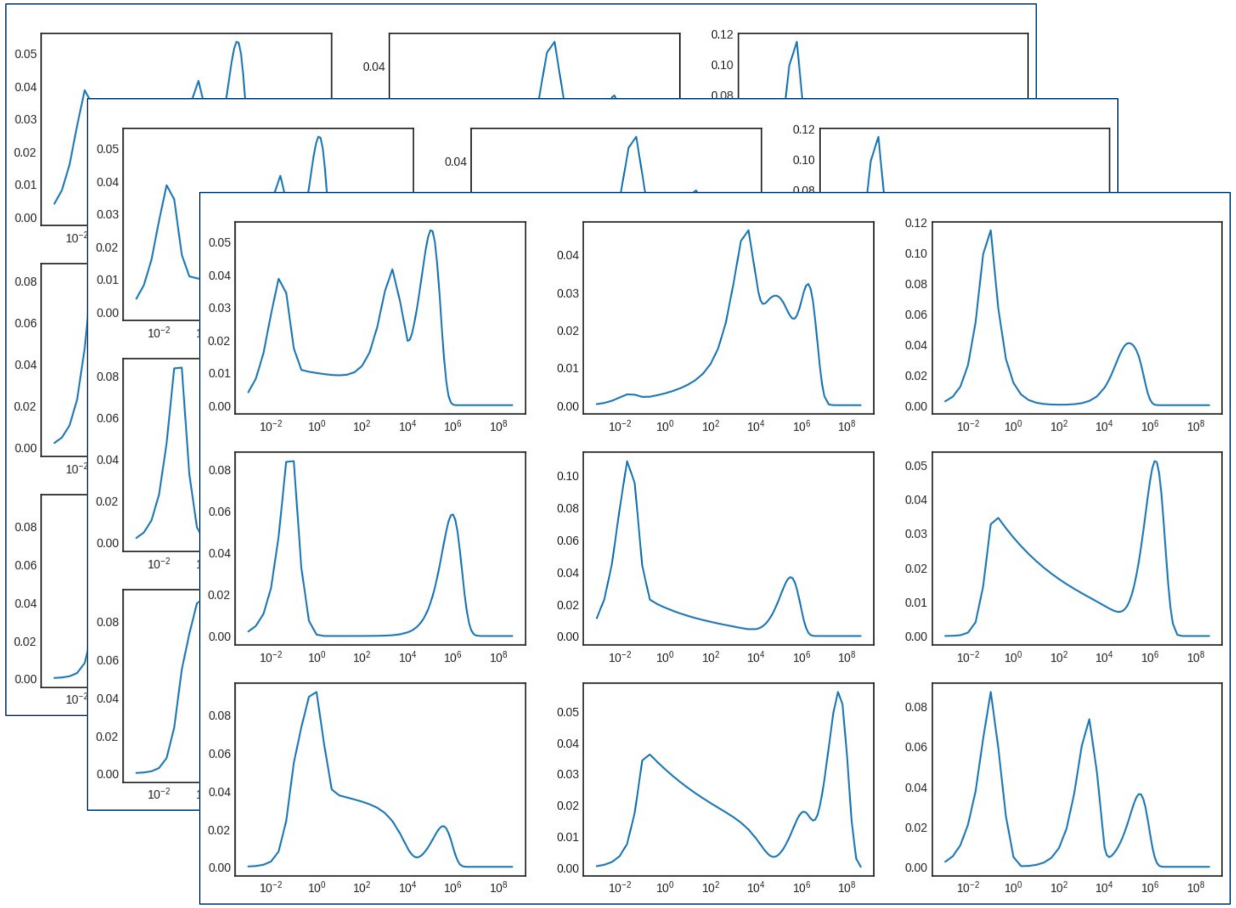


DOI: [10.9734/air/2020/v21i930238](10.9734/air/2020/v21i930238)

Dataset (80/20% train/test)

1. **IAEA dataset* (251 spectra)**
    1. Input features:
        a. Numerical: 7 measurements
        b. Category: one-hot encoding, 8 categories
    2. Output target: Spectre values for 60 energy bins (10^{-8} – $3.98 \cdot 10^{2}$ MeV, logscale)

2. **Synthetic based on FRUIT** method (10^{4} spectra)**
    1. Input features: numerical, 7 measurements
    2. Output target: Spectre values for 60 energy bins (10^{-8} — $3.98 \cdot 10^{2}$ MeV, logscale)



**Compendium of Neutron Spectra and Detector Response for Radiation Protection Purposes: Technical Report Series. Vienna: IAEA, 2001. No. 403.*

** Frascati Unfolding Interactive Tool, doi: 10.1016/J.NIMA.2007.07.033

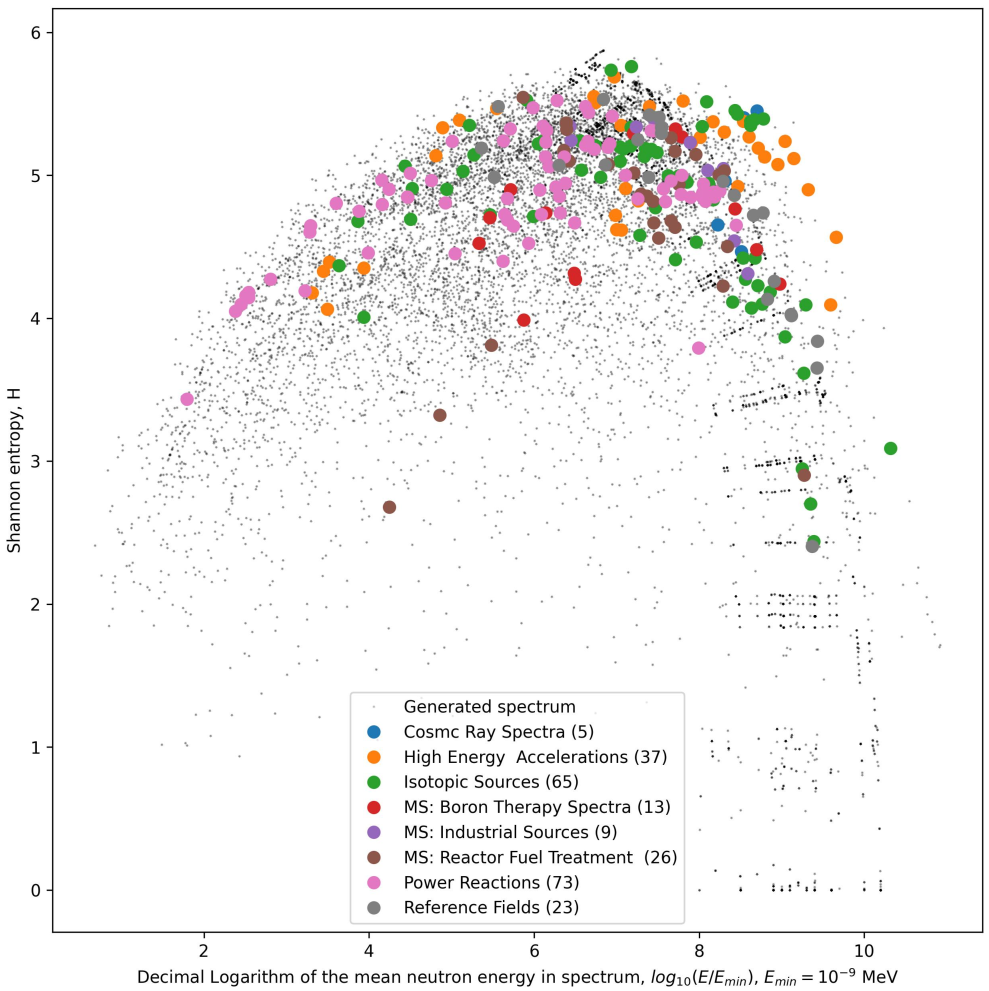
Generated and real spectra

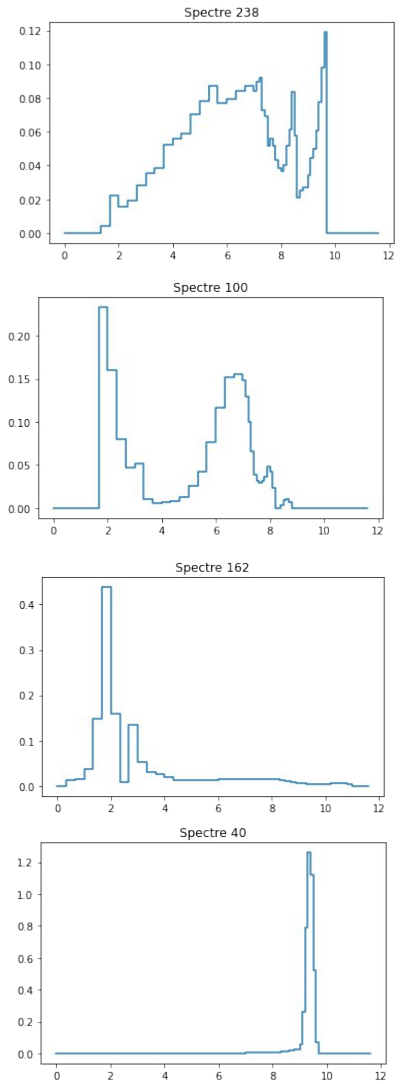
Spectrum representation in FRUIT method:

$$\varphi(E) = P_{th}\varphi_{th}(E) + P_e\varphi_e(E) + P_f\varphi_f(E) + P_{hi}\varphi_{hi}(E),$$

$$P_{th} + P_e + P_f + P_{hi} = 1$$

- φ_{th} — Maxwell thermal
- φ_e — epitermal
- φ_f — fast
- φ_{hi} — high-energy





ML model for IAEA dataset

$M = 7$
$K = 8$
$S = 60$

Q_1

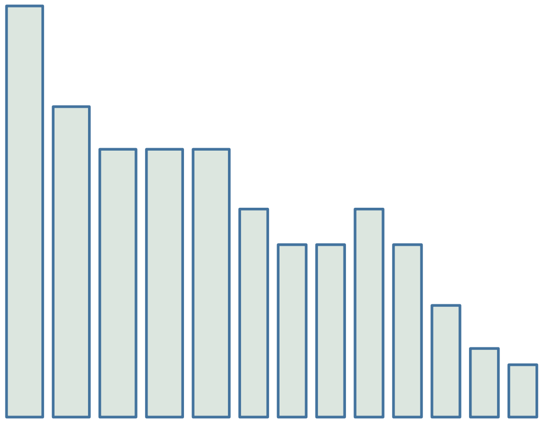
Q_2

Q_M

Category_1

Category_2

Category_k

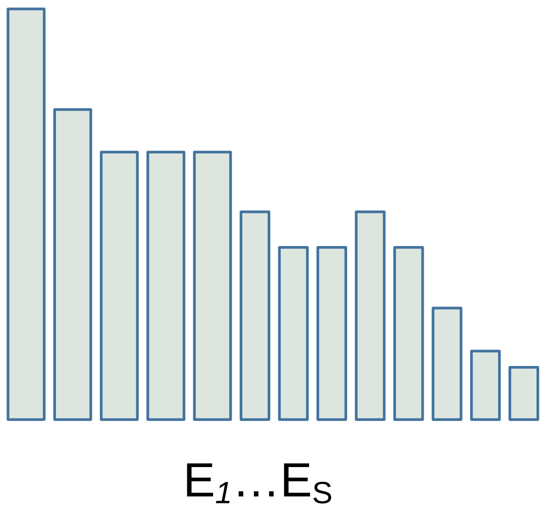


$E_1 \dots E_S$

$$Q_j \approx \sum_{i=1}^{S} K_{ji} \Phi(\mathrm{E}_i) \Delta E, \quad j = 1, \dots, M$$

ML model for synthetic spectra

$M = 7$
$S = 60$

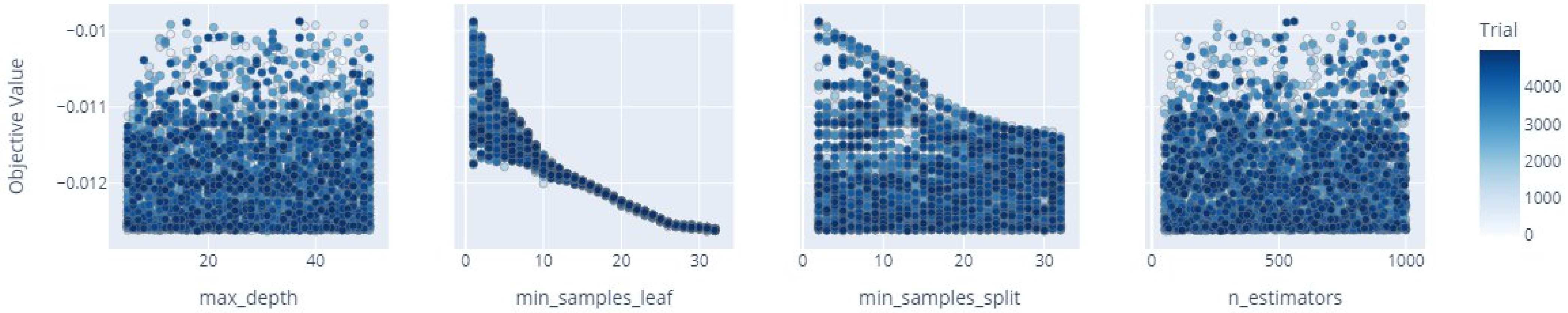


$$Q_j \approx \sum_{i=1}^{S} K_{ji} \Phi(\mathrm{E}_i) \Delta E, \quad j = 1,\ldots,M$$

Optimization of the algorithm

The criterion for measuring the quality of separation is the root mean square error.

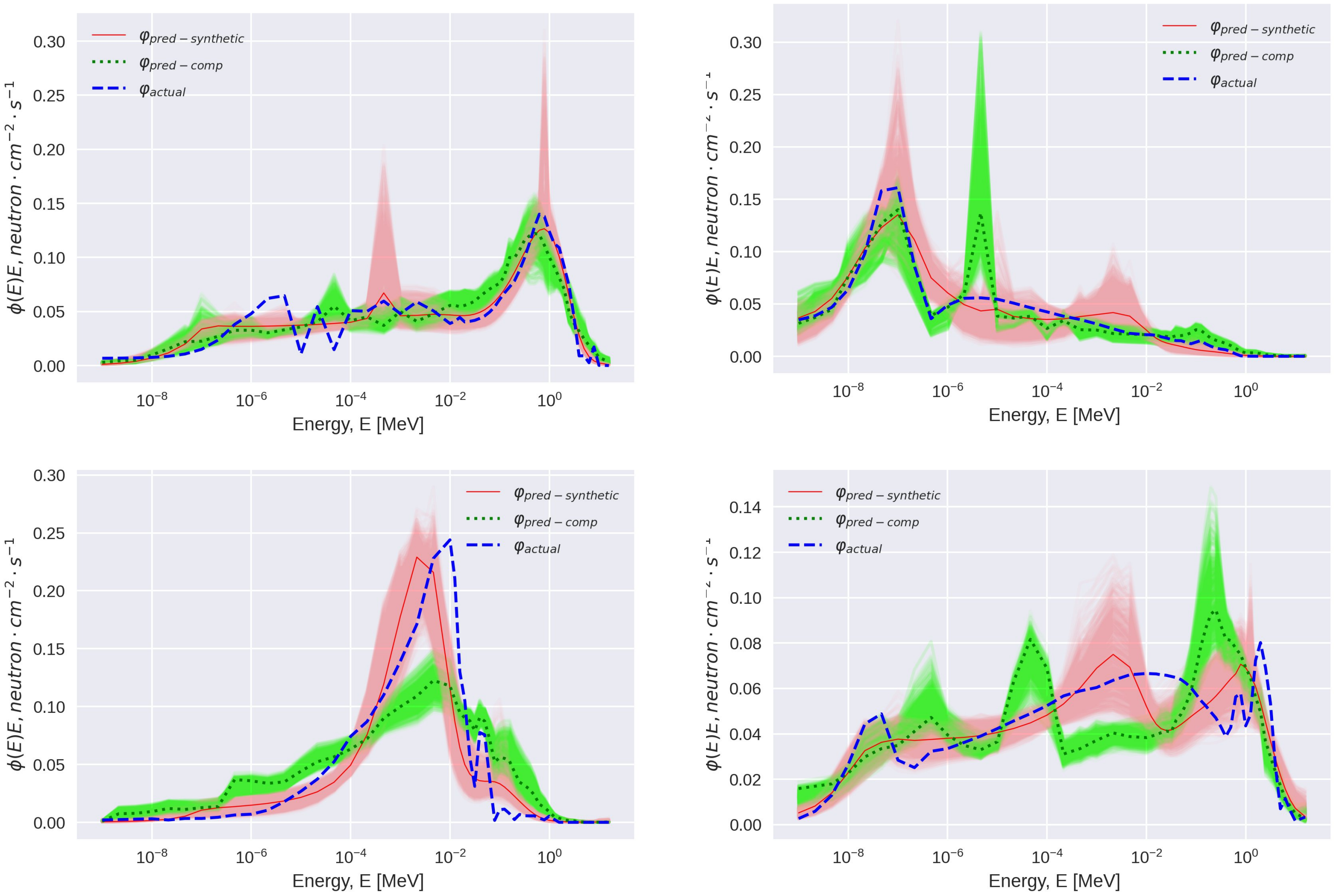
Optimization criterion: negative mean maximum error.

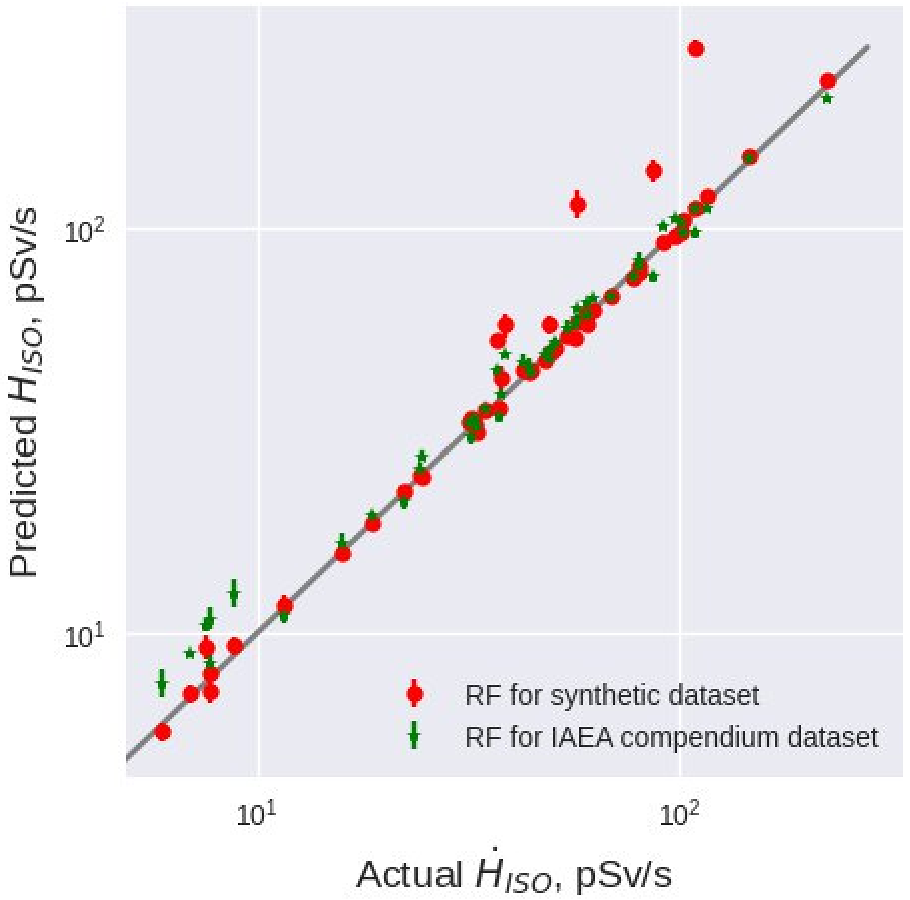


- the number of trees in the forest (n_estimators) = 169;
- the maximum tree depth (max_depth) = 39;
- the minimum number of samples required to split an internal node (min_samples_split) = 2;
- the minimum number of samples required to be at a leaf node (Min_samples_leaf) = 1.

Results

- IAEA dataset R^2 score = 0.57 MAE = $1.9 \cdot 10^{-2}$ MSE = $1.3 \cdot 10^{-3}$
- synthetic dataset R^2 score = 0.67 MAE = $1.4 \cdot 10^{-2}$ MSE = $1.6 \cdot 10^{-3}$

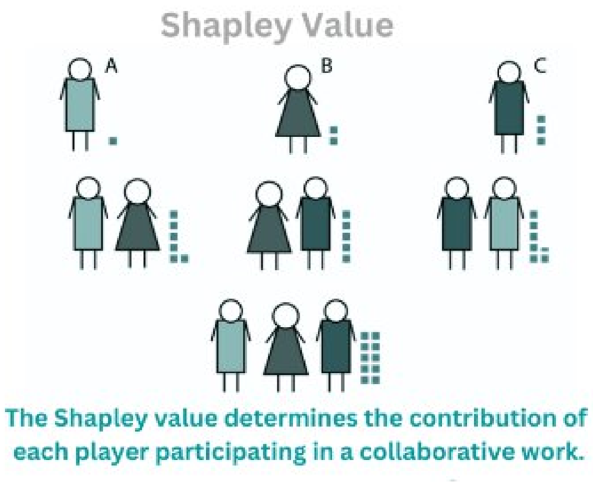




uncertainty of spectrum unfolding for $\xi = 5\%$ error in measurements

SHAP (Shapley Additive Explanation)

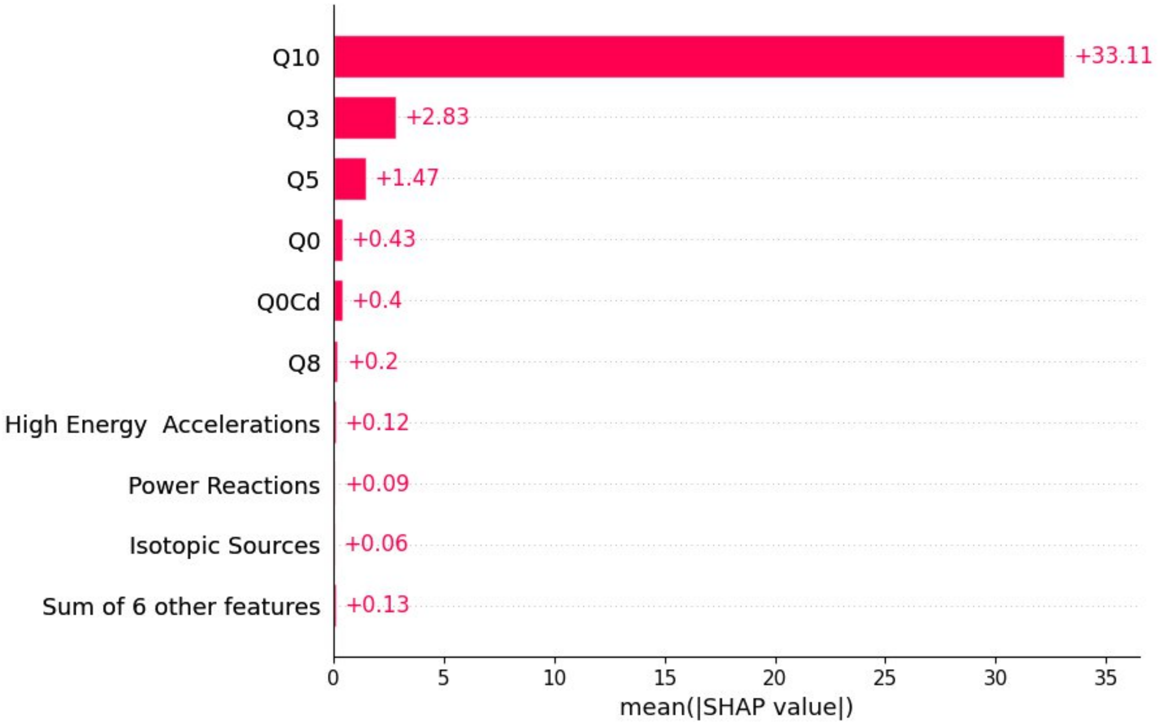
A method for explaining individual forecasts. SHAP is based on the game of theoretically optimal Shapley values. The technical definition of the Shapley value is “the average marginal contribution of a feature value to all possible coalitions...”

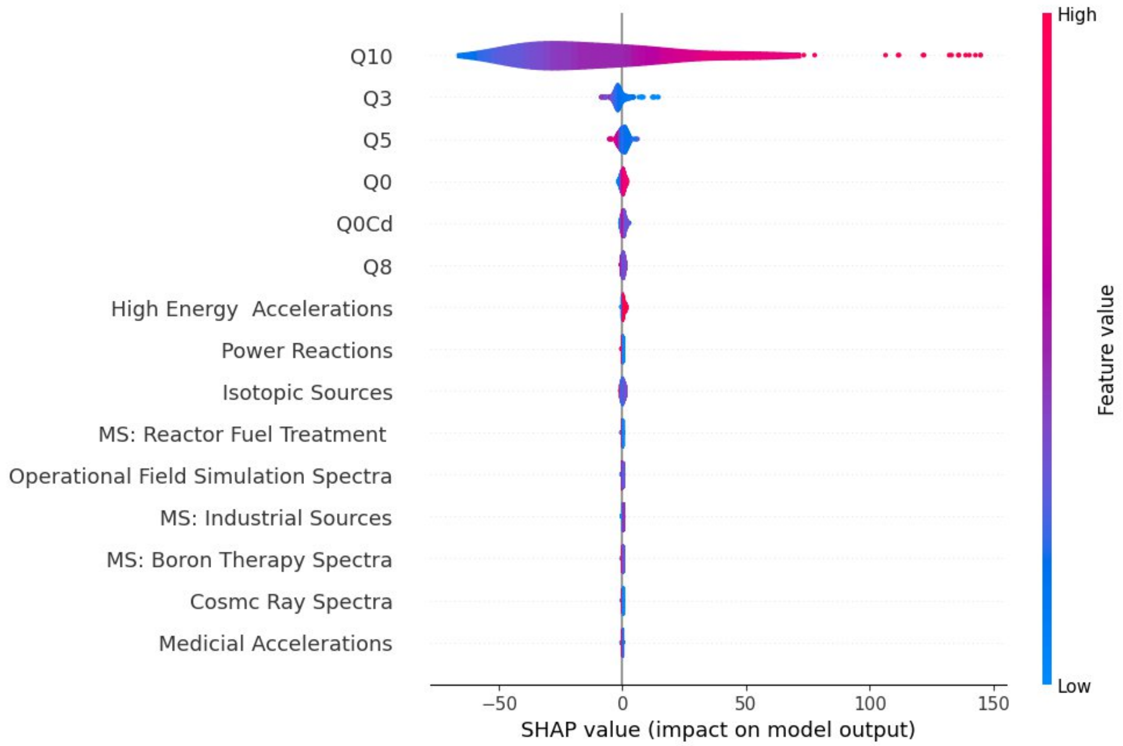


https://shorturl.at/GBgDZ

SHAP (Shapley Additive Explanation)

Impact of the measurement with moderator sphere to the dose rate assessment





Auto ML

AutoML solutions automatically develop ML-based models. We used *lightautoml** package.

Automatic fine-tuning of the model hyperparameters and blending of different models:

1. Linear regression with L2 regularization
2. LightGBM
3. CatBoost

LOSS = MAE
METRIC = MSE

Final prediction for new objects =
0.22569 * (5 averaged models Lvl_0_Pipe_2_Mod_0_CatBoost) +
0.77431 * (5 averaged models Lvl_0_Pipe_2_Mod_1_Tuned_CatBoost)

Prediction metrics:
MAE: $4.2 \cdot 10^{-3}$
MSE: $2.0 \cdot 10^{-4}$
R^2 score: 0.80

*Vakhrushev, Anton, et al. "Lightautoml: Automl solution for a large financial services ecosystem", *arXiv preprint arXiv:2109.01528* (2021).

Discussion

- Development of a training set using Monte Carlo simulations (Geant4).
- Improving the set of moderator spheres in the **high-energy region** (adding composite spheres with lead and other materials).
- Other metrics metrics of the discrepancy between the true and unfolded distributions: Kolmogorov-Smirnov statistic, maximum mean discrepancy.
- Combine the algorithm with other methods to improve accuracy

Conclusions

- This study presents a machine learning-based approach for neutron spectrum unfolding and effective dose rate estimation using Bonner spectrometer measurements.
- Comparison of two Random Forest Regression models trained on a small real spectres dataset and on a large synthetic dataset is presented. Both models showed strong correlation between actual and unfolded spectra and in dose rate assessment.
- The SHAP method enabled the optimization of BSS moderator sphere selection, it is showed that measurements with the 10” moderator sphere make are the most important to the dose assessment.
- The proposed method could be used for improving radiation protection in high-energy neutron fields.

Authors acknowledge the support of the JINR Multifunctional Information and Computing Complex for providing computational resources, http://hlit.jinr.ru/.
The research was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation (theme No. 124092700007-4).

Thank you!

kchizhov@jinr.ru

Publications:

1. Chizhov, K., L. Beskrovnaya, and A. Chizhov. "Neutron Spectra Unfolding from Bonner Spectrometer Readings by the Regularization Method Using the Legendre Polynomials." *Physics of Particles and Nuclei* 55.3 (2024): 532-534.
2. Chizhov K, Chizhov A. Dose assessment of personnel neutron irradiation on high-energy accelerators using a multi-sphere Bonner spectrometer. *Mathematical Modeling* 7 (2), 63-64, 2023.