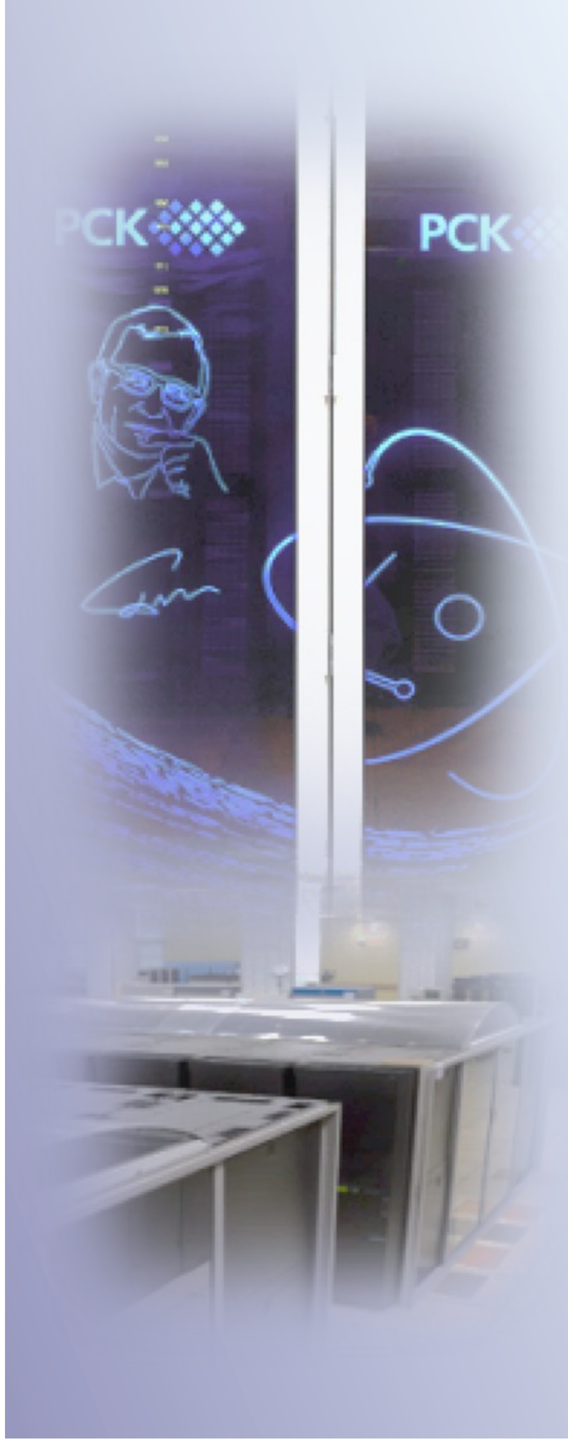
Automated machine learning spectrum unfolding for neutron spectrometry with Bonner spheres

Chizhov K.A.[1,2]

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

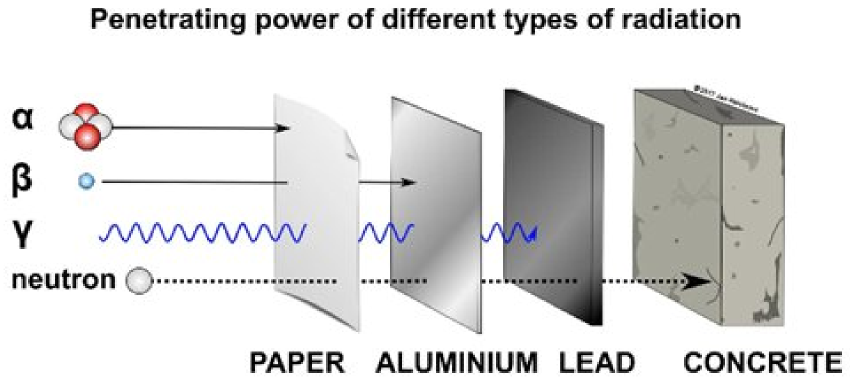
[2]University “Dubna”, Dubna, Russia.

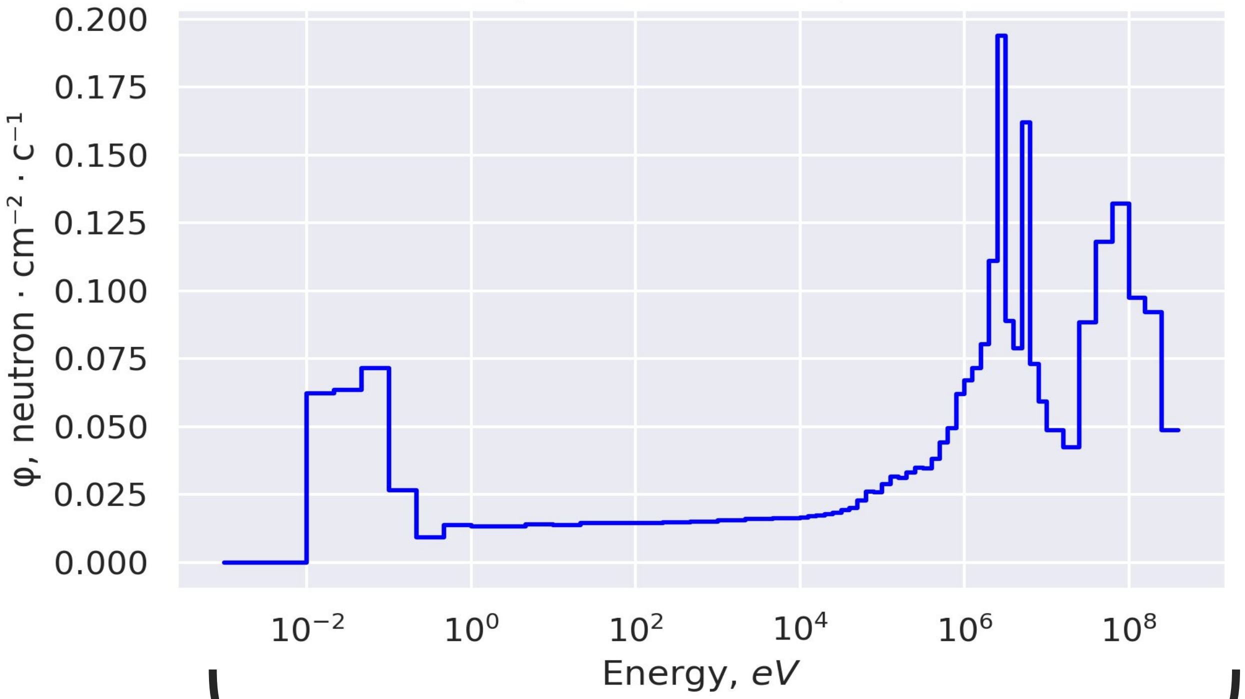


The 11th International Conference "Distributed Computing and Grid-technologies in Science and Education"
July, 7-11, 2025 JINR, Dubna, Russia.

Introduction

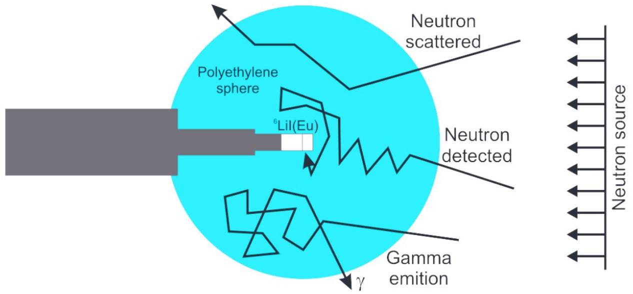
Radiation fields behind the protective shields of nuclear physics facilities (particle accelerators, nuclear reactors) are formed mainly by **neutrons** of a wide energy spectrum.





Neutron energy range

- Nuclear power plants, E ≈ 10^{-9} – 20.5 MeV
- Particle accelerators, E ≈ 10^{-9} – 10^{3} MeV
- Cosmic radiation, E ≈ 10^{-9} – 10^{4} MeV



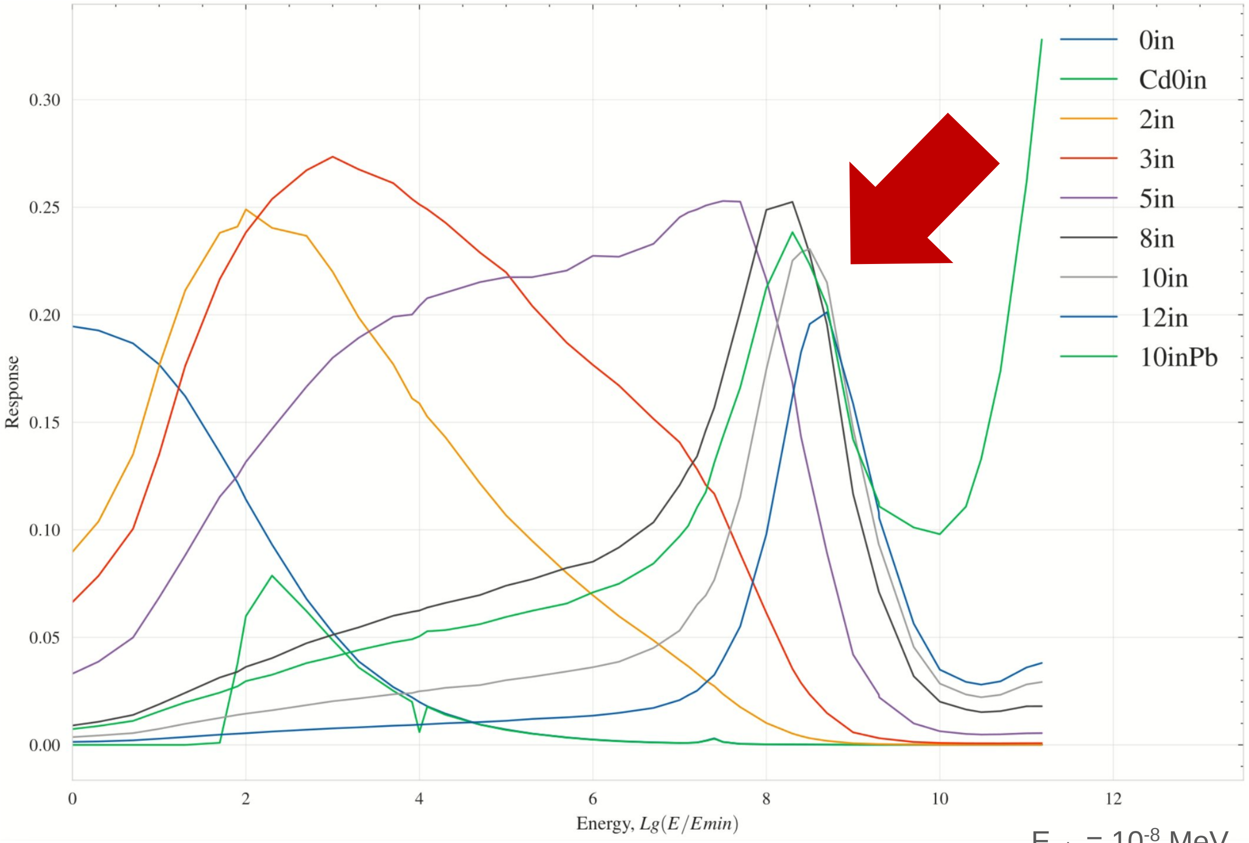
http://dx.doi.org/10.5772/51274

Personal dosimeter
0.025 eV - 20 MeV
https://www.doza.ru

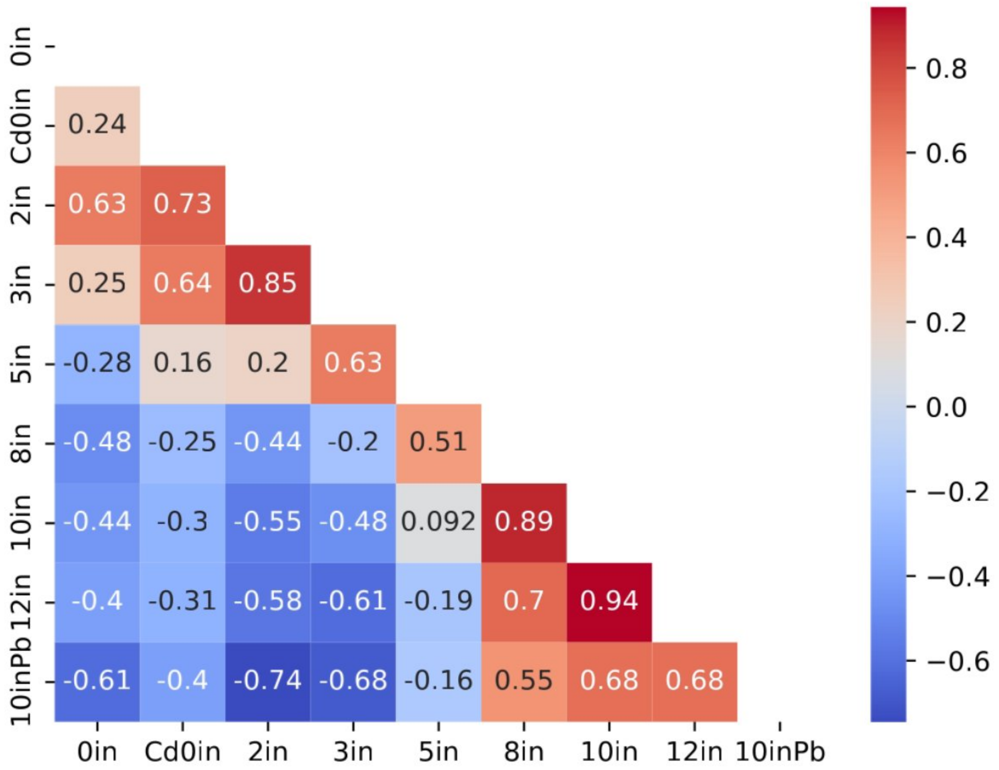
Bonner
Multi-sphere
Spectrometer

https://atomtex.com/

Response functions of moderator spheres



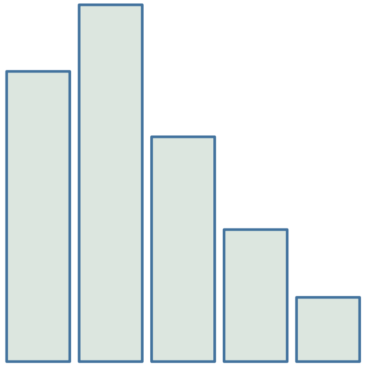
$E_{min} = 10^{-8}$ MeV



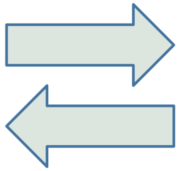
Martinkovic, J., and G. N. Timoshenko. *Calculation of multisphere neutron spectrometer response functions in energy range up to 20 MeV*. No. JINR-R--16-2005-105. Division of Radiation and Radiobiological Research, 2005.

Direct and inverse problems of spectrum unfolding

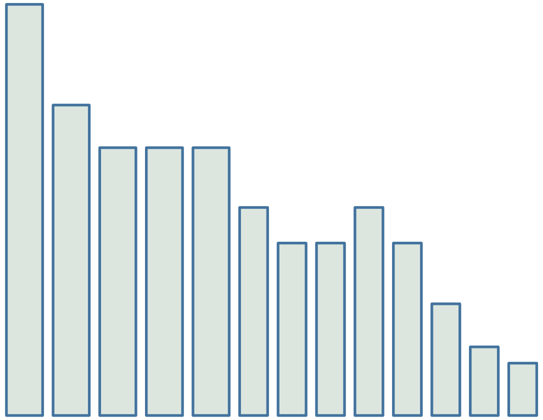
- Direct task: to obtain “effective” readings of the Bonner spectrometer (BSS) based on the given spectra.
- Inverse task: to obtain the initial spectra based on the measurements.



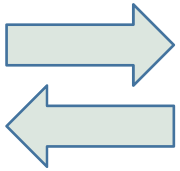
$Q_1 \ldots Q_M$

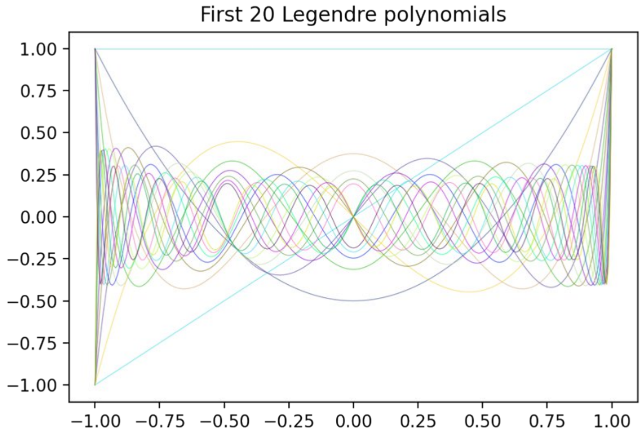


$M < S$
$M < N$



$E_1 \ldots E_S$





$C_1 \ldots C_N$

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

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

$$P_n(x) = \frac{1}{2^n n!} \cdot \frac{d^n}{dx^n} \left(x^2 - 1\right)^n$$

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

- Q_j is the BSS reading for the j-th sphere,
- ξ is the error for measurement,
- $\varphi(E)$ is the neutron spectrum,
- $K_j(E)$ is the response function of the detector to neutrons,
- M is the number of spheres used,
- The integration limits $E_{\min}$ and $E_{\max}$.

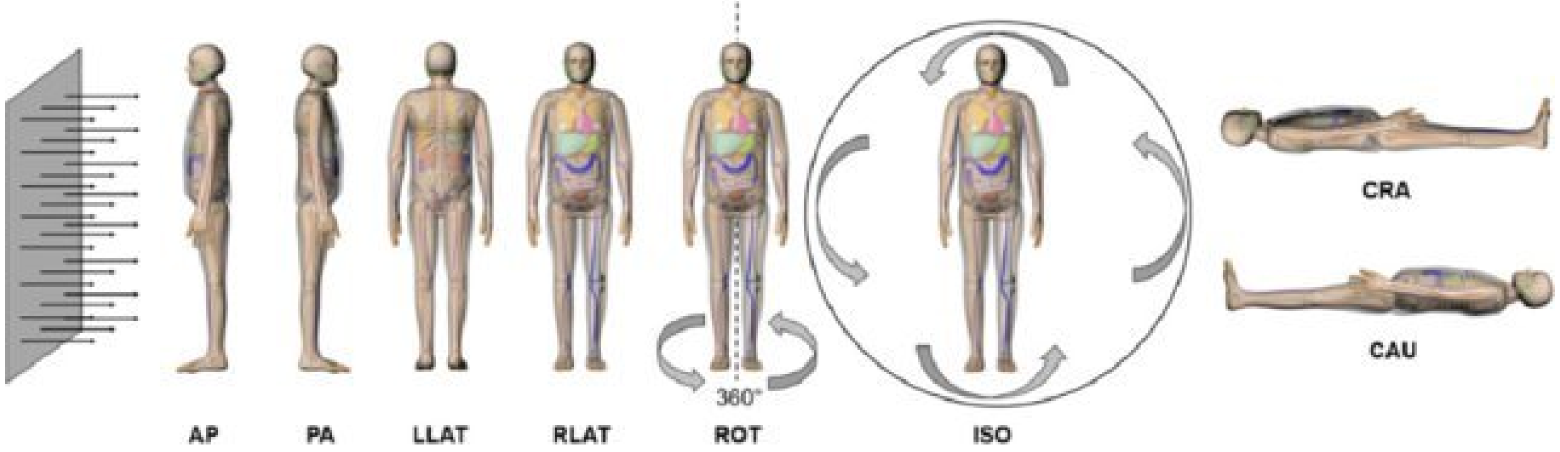
Dose estimation for spectrum

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

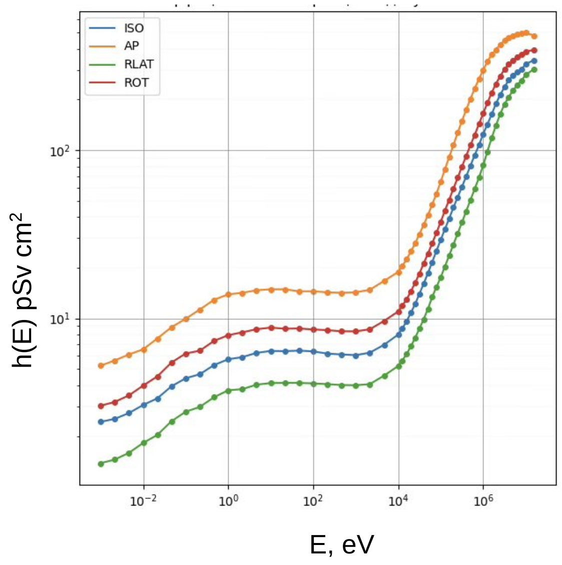
$$\dot{H} = \int_{Emin}^{Emax} \varphi(E)\cdot h(E)dE$$

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

h(w) – corresponding conversion factor [pSv·cm^2] for monoenergetic particles in different irradiation geometries, ICRP Publication 116[1].



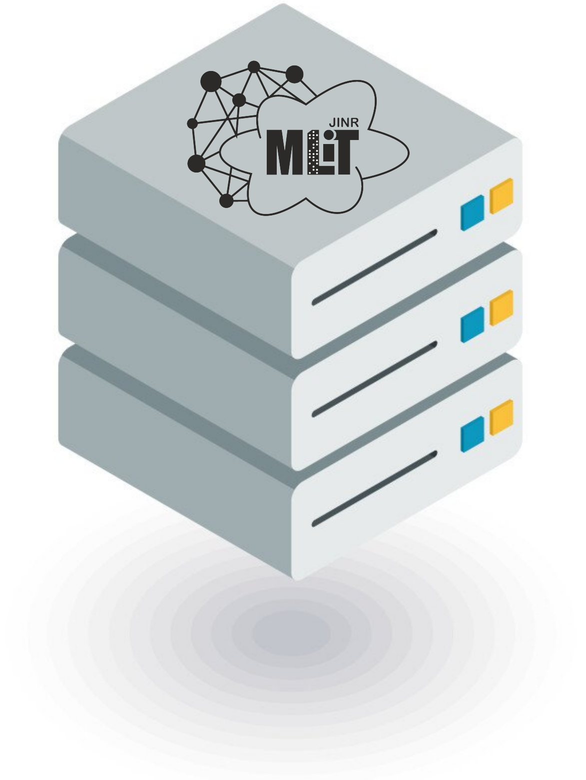
https://dceg.cancer.gov/tools/radiation-dosimetry-tools/phantoms-library



1. 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.

ML method & implementation

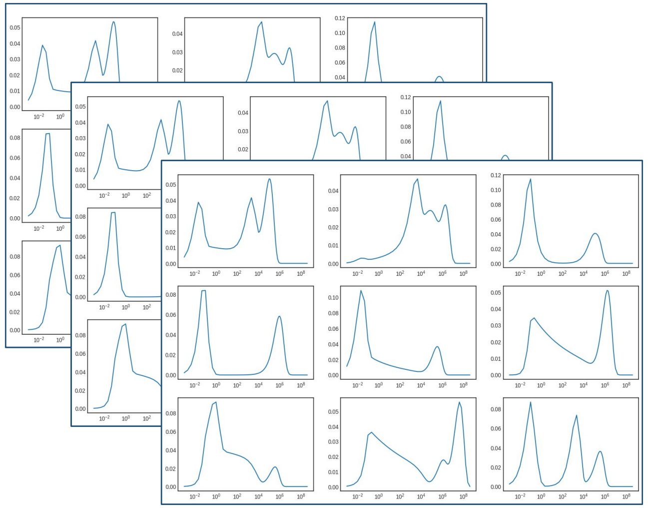
- Calculations on the multifunctional information and computing complex of the Joint Institute for Nuclear Research (JINR) Information Technology Laboratory.
- Automated machine learning (AutoML) frameworks: Fedot, LightAutoML.



1. http://hlit.jinr.ru/
2. https://github.com/aimclub/FEDOT
3. https://github.com/sb-ai-lab/LightAutoML

Dataset

- **Synthetic based on FRUIT[1] method ($5 \cdot 10^5$ spectra) for training**
  1. Input features: numerical: 10 measurements.
  2. Output target: Spectre values for 60 energy bins (10^{-9} - $6.3 \cdot 10^2$ MeV, log scale).
  3. 80% — training, 20% — validation.



- **Spectra from the literature (375 spectra) for testing**

IAEA dataset[2] (251 spectra) + 124 spectra from open access papers (manually digitized by our group).

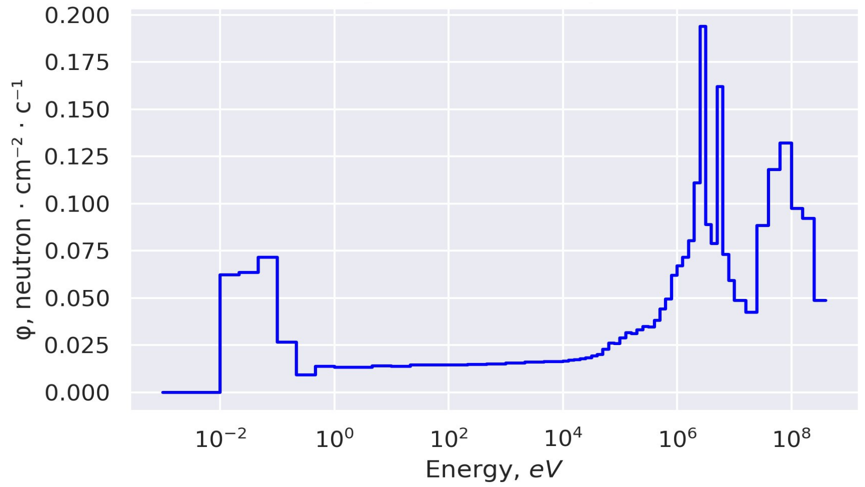
1. Input features: numerical: 10 measurements.
2. Output target: Spectre values for 60 energy bins (10^{-9} — $6.3 \cdot 10^2$ MeV, log scale).

Monte-Carlo simulation, regularization (Tikhonov, statistical, ...), Maximum entropy principle, Maximum likehood, Genetic, Iterative, Parametric, Bayesian, ...[3]

1. Frascati Unfolding Interactive Tool, doi: 10.1016/J.NIMA.2007.07.033
2. *Compendium of Neutron Spectra and Detector Response for Radiation Protection Purposes: Technical Report Series. Vienna: IAEA, 2001. No. 403.*
3. 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.

Synthetic dataset

Model	Thermal $\phi_{\text{th}}\ (E,...)$	Epithermal $\phi_e\ (E,...)$	Fast $\phi_f\ (E,...)$	High Energy $\phi_{\text{hi}}\ (E,...)$
Fission	$\left(\frac{E}{T_0^2}\right)e^{-E/T_0}$	$\left[1-e^{-(E/E_d)^2}\right]E^{b-1}e^{-E/\beta'}$	$E^{\alpha}e^{-E/\beta}$	0
Evaporation	↓	↓	$\left(\frac{E}{T_{ev}^2}\right)e^{-E/T_{ev}}$	0
Gaussian	↓	↓	$exp\left(-\frac{1}{2}\left(\frac{E-E_m}{\sigma E_m}\right)\right)$	0
High Energy	↓	↓	$\left(\frac{E}{T_{ev}^2}\right)e^{-E/T_{ev}}$	$\left(\frac{E}{T_{hi}^2}\right)e^{-E/T_{hi}}$



The tunable parameters for each model are:

Fission: $P_{th},P_e,P_f,b,\beta',\alpha,\beta$Pth,Pe,Pf,b,β′,α,β
Evaporation: $P_{th},P_e,P_f,b,\beta',T_{ev}$Pth,Pe,Pf,b,β′,Tev
Gaussian: $P_{th},P_e,P_f,b,\beta',E_m,\sigma$Pth,Pe,Pf,b,β′,Em,σ
High Energy: $P_{th},P_e,P_f,P_{hi},b,\beta',T_{ev},T_h$

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

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

Type	Energy	Velocity (m/s)	Wavelength (nm)	Temperature (K)
ultracold	< 0.2 μeV	< 6	> 64	< 0.002
very cold	$0.2 \le E < 50$ μeV	$6 \le v < 100$	4< $\lambda \le 64$	0.002 ≤ T < 0.6
cold	$0.05 < E \le 3$ meV	$100 < v \le 760$	0.5 ≤ λ <4	0.6 < T ≤ 34
thermal	$0.003 \le E < 0.04$ eV	$760 < v \le 2760$	0.14≤ λ <0.5	34 < T ≤ 0.46
epithermal	0.04 eV< $E \le 100$ keV	$2760 < v \le 4.3 \cdot 10^6$	$92 \cdot 10^{-6} \le \lambda < 0.14$	$(0.46 < T \le 1.2) \cdot 10^6$
intermediate	$100 < E \le 200$ keV	$(4.3 < v \le 6.2) \cdot 10^6$	$(64 \le \lambda < 92) \cdot 10^{-6}$	$(1.2 < T \le 2.3) \cdot 10^6$
fast	$0.2 < E \le 10$ MeV	$(6.2 < v \le 43) \cdot 10^6$	$(9.2 \le \lambda < 64) \cdot 10^{-6}$	$(0.002 < T \le 120) \cdot 10^9$
high-energy	> 10 MeV	$> 4.3 \cdot 10^7$	$< 9.2 \cdot 10^{-6}$	$> 120 \cdot 10^9$

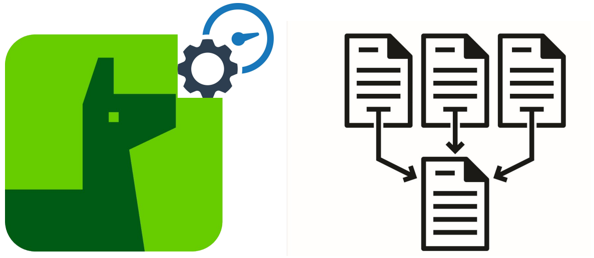
1. *McGreivy J., Manfredi J.J., Siefman D. Data Augmentation for Neutron Spectrum Unfolding with Neural Networks. Journal of Nuclear Engineering 4(1)): 77–95. 2023. (For neutron energy range from* 0.001 eV to 15.8 MeV*)*
2. Frascati Unfolding Interactive Tool, doi: 10.1016/J.NIMA.2007.07.033
3. Nico S., Snow W. M. Fundamental Neutron Physics. Annu. Rev. Nuc.Part. Sci., 55:27–69, 2005.

LightAutoML (LAMA)

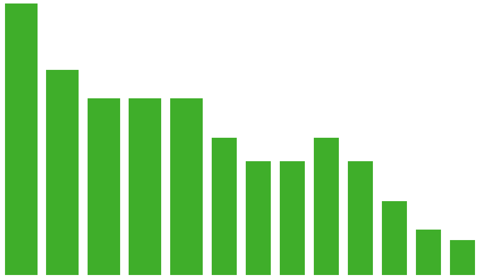
AutoML solutions automatically develop ML-based models. LightAutoML provides automatic fine-tuning of the model hyperparameters and blending of different models: Linear regression with L2 regularization, LightGBM, CatBoost, Random forest (in multi- regression mode).

SLAMA – *distributed* version of the LightAutoML library (LAMA) for *Spark 3+* for processing large datasets.

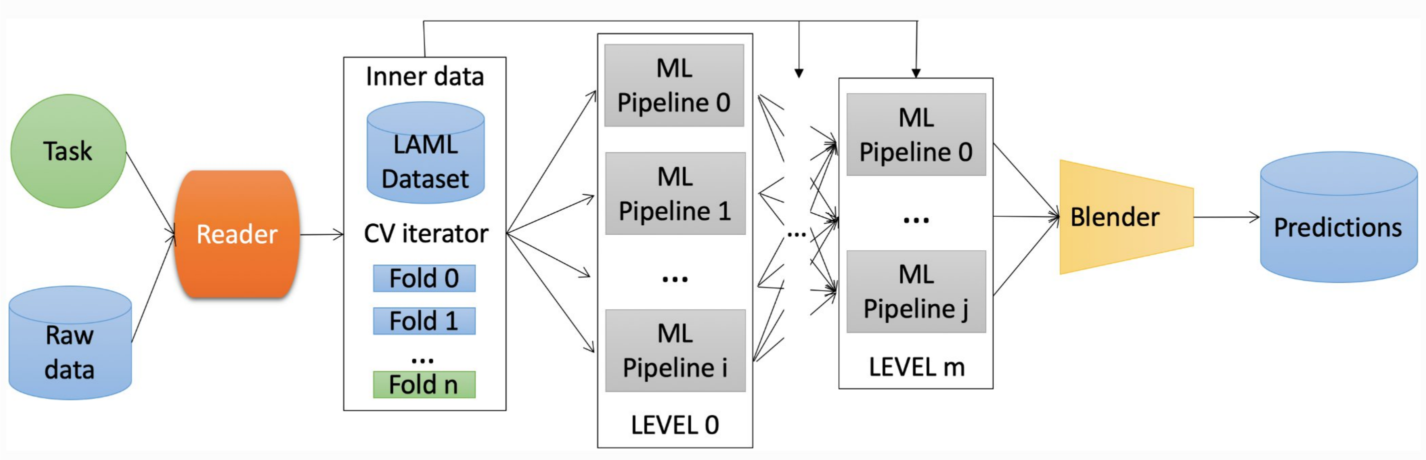
Measurements



Composite model

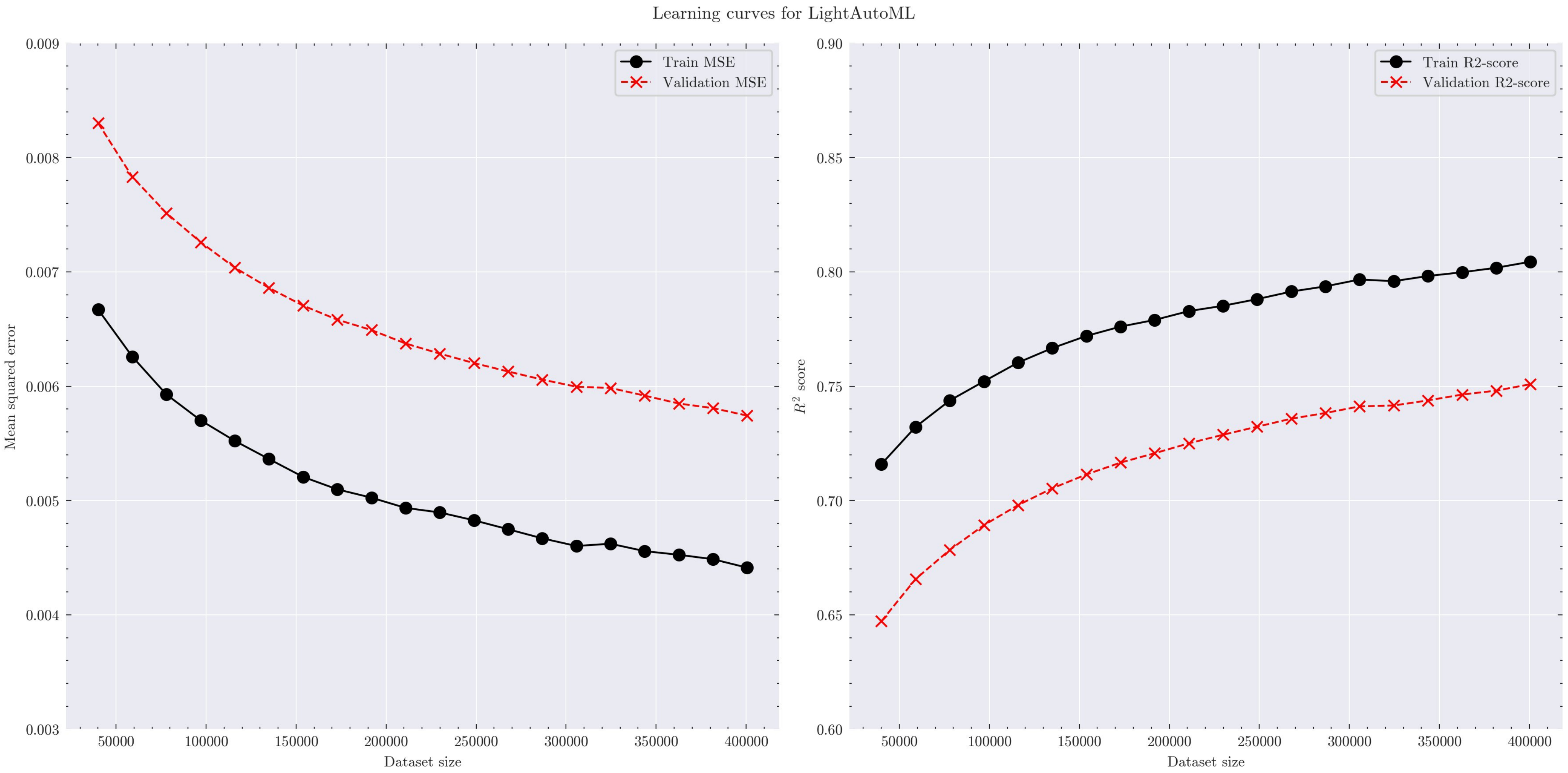


$E_1 \dots E_S$



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

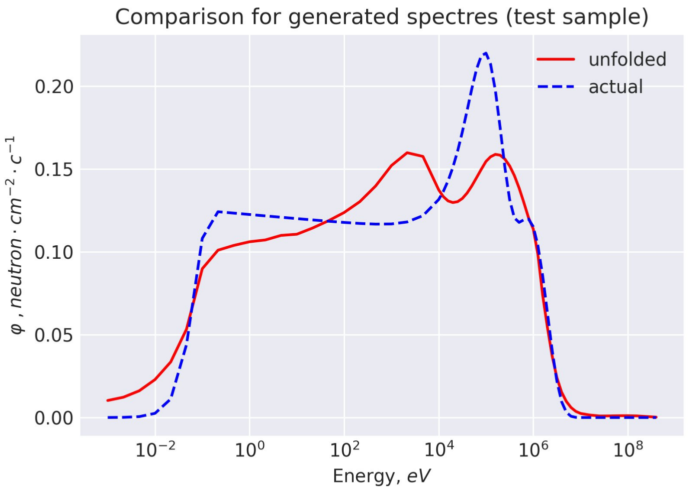
LightAutoML. Learning curve

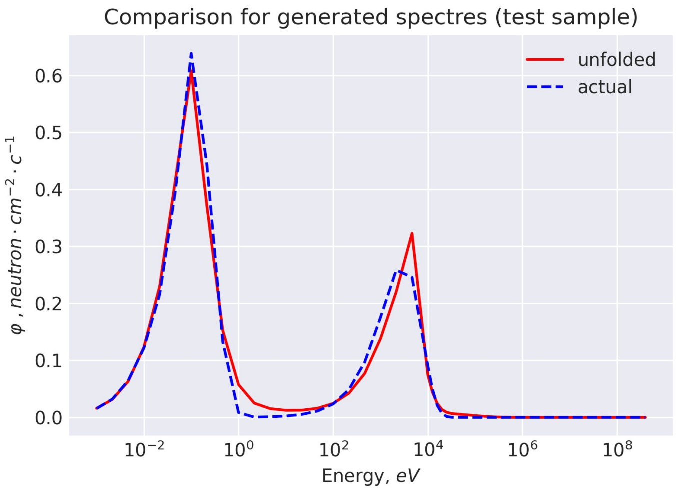


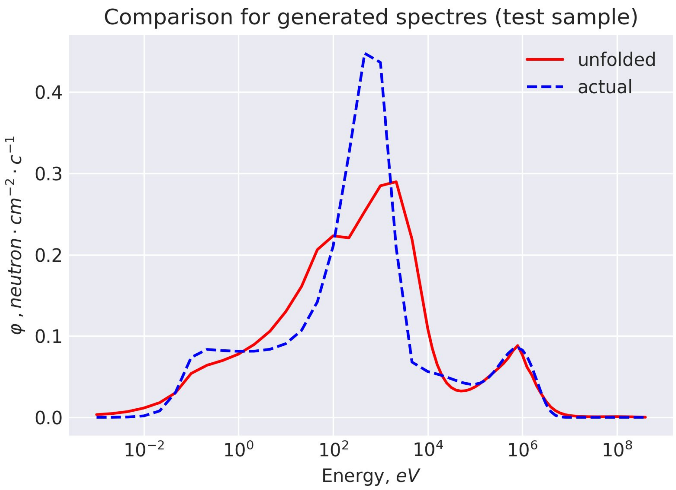
Loss = MAE (mean absolute error),
Metric = MSE.

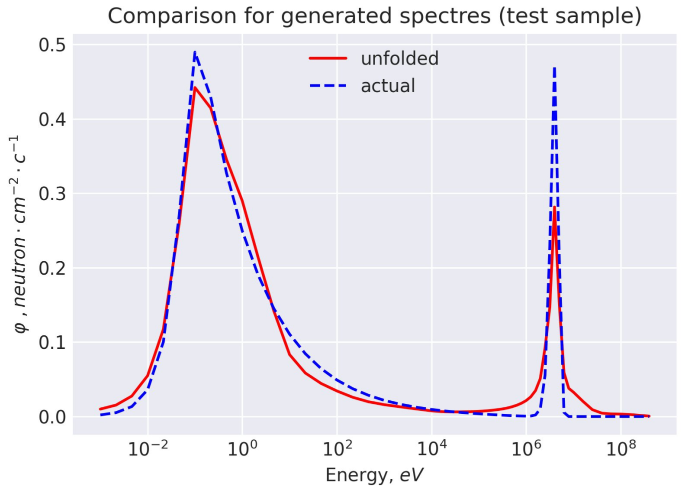
Final model: Random forest.

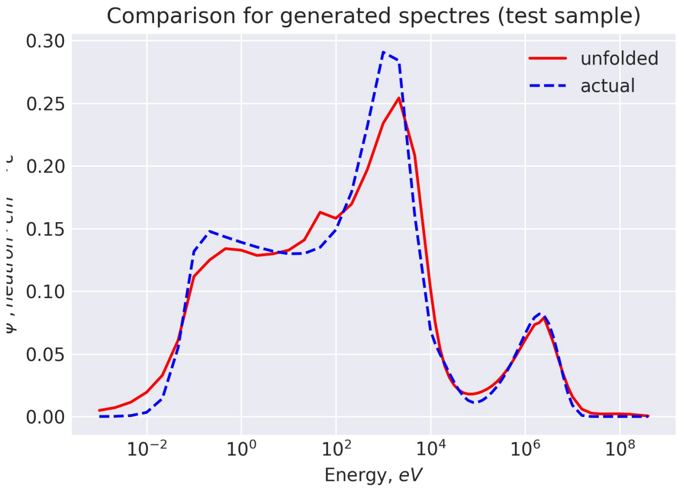
LightAutoML results. Comparison for generated spectra

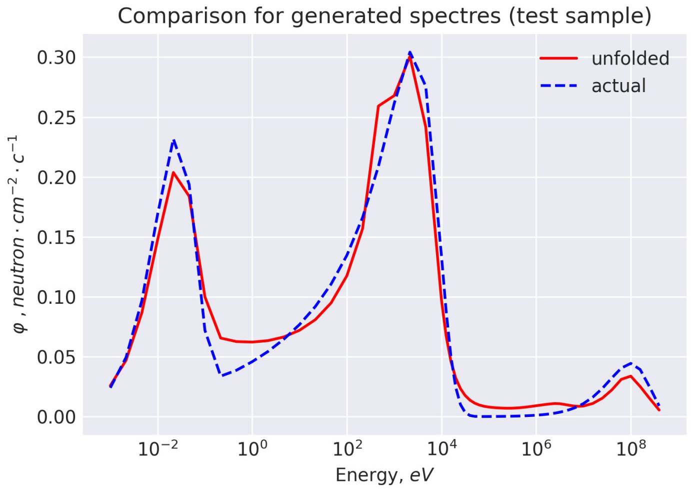




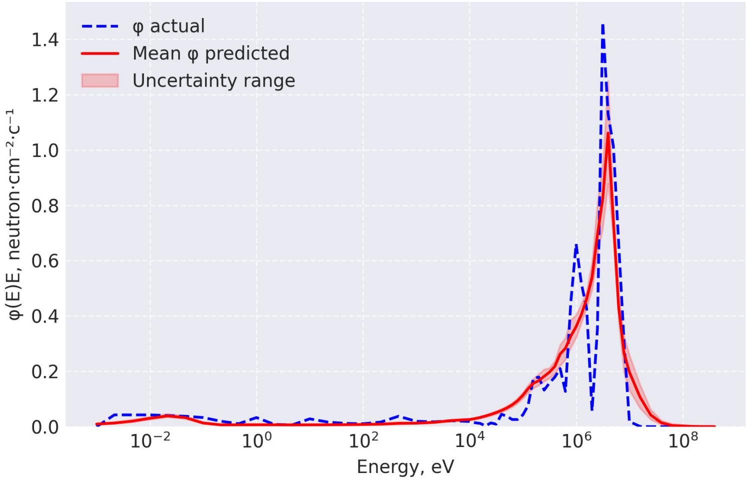




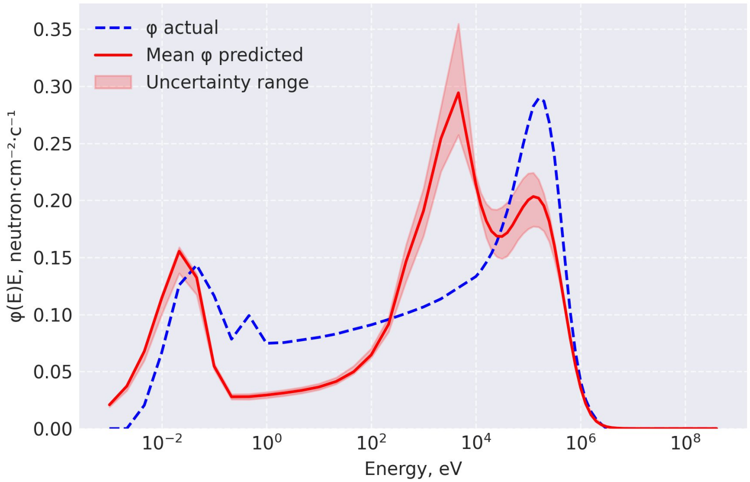




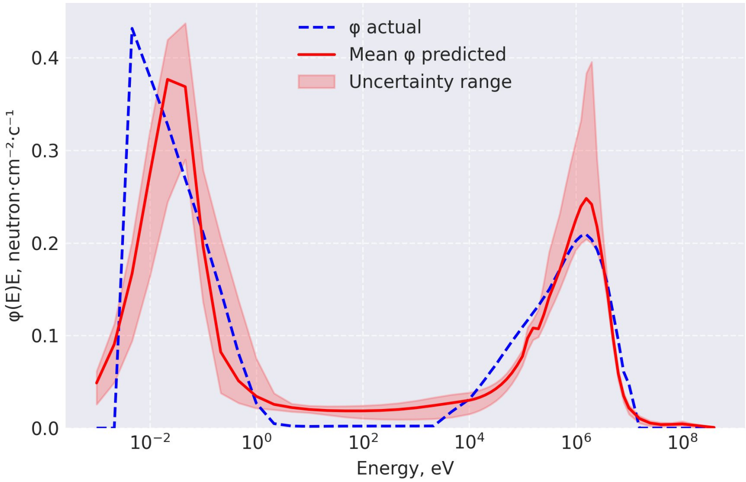
LightAutoML results. Spectra comparison for 375 real spectra



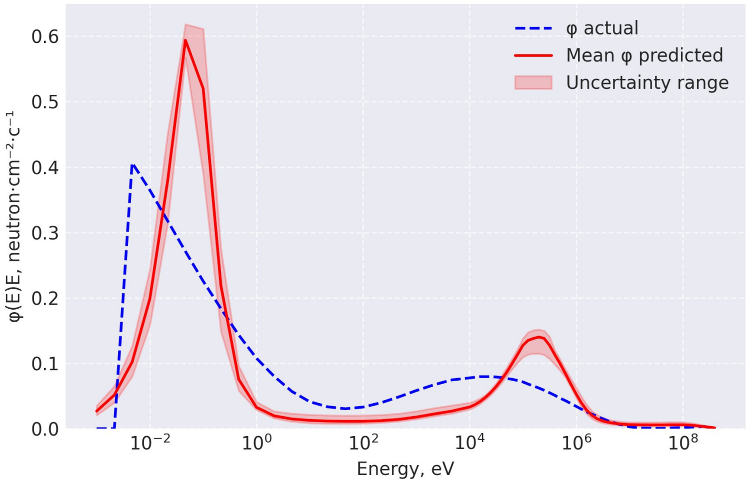
Spectra №0



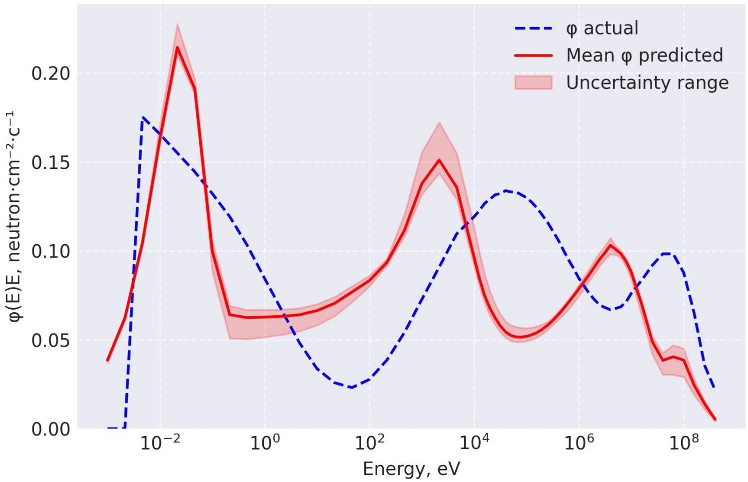
Spectra №102



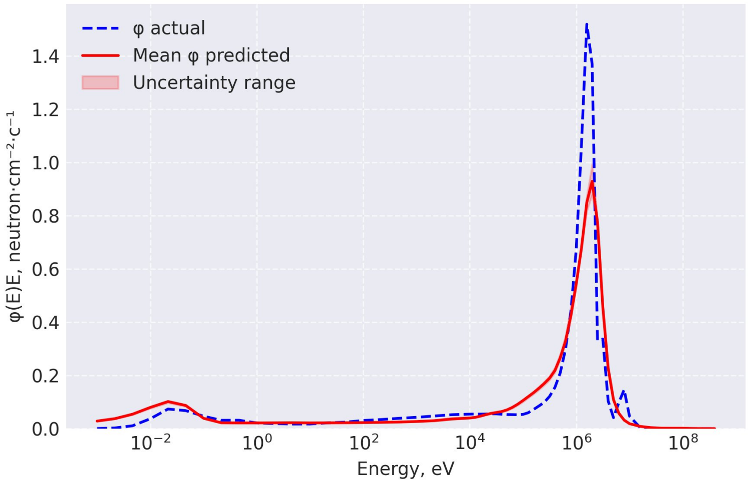
Spectra №83



Spectra №67 (soft field)



Spectra №68 (hard field)

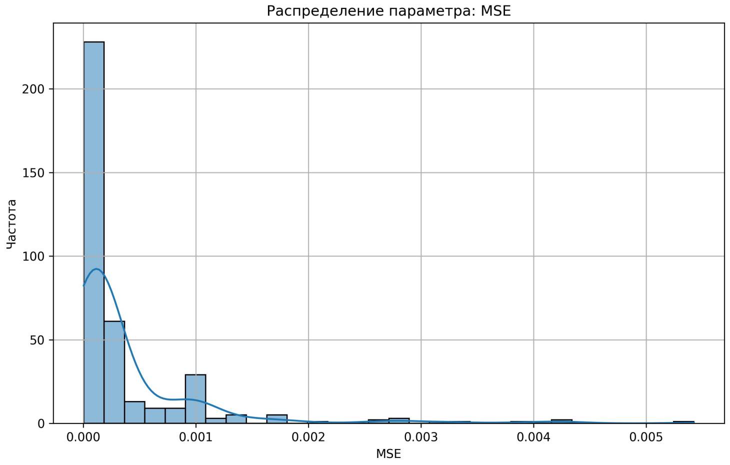


Spectra №226

The uncertainty of the spectra unfolding was estimated using the Monte Carlo method, in which random perturbations (1%) were introduced into the input data.

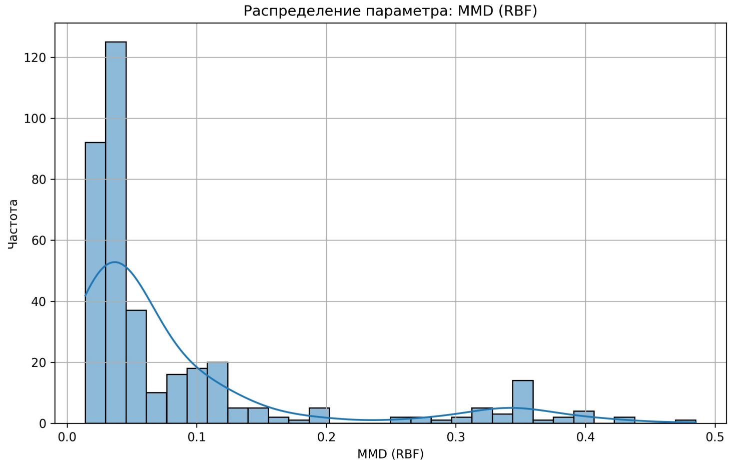
Timeout = 720 minutes

LightAutoML results. Spectra comparison for 375 real cases



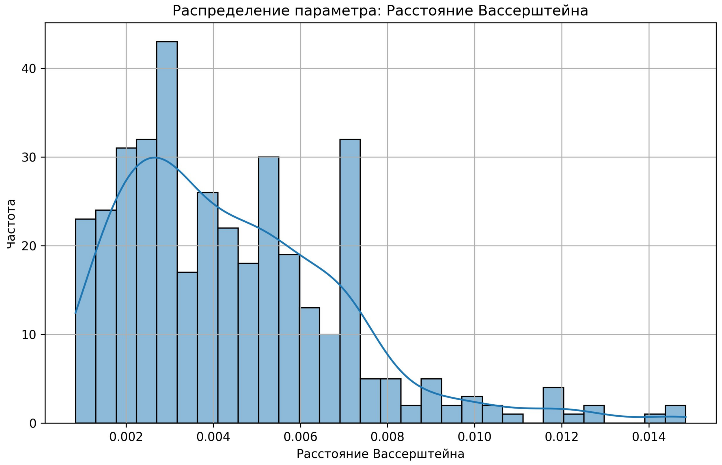
MSE

Mean squared error



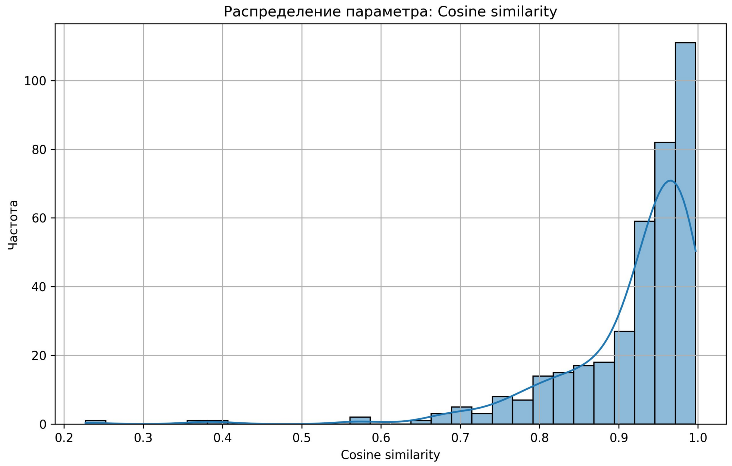
MMD

Maximum mean discrepancy

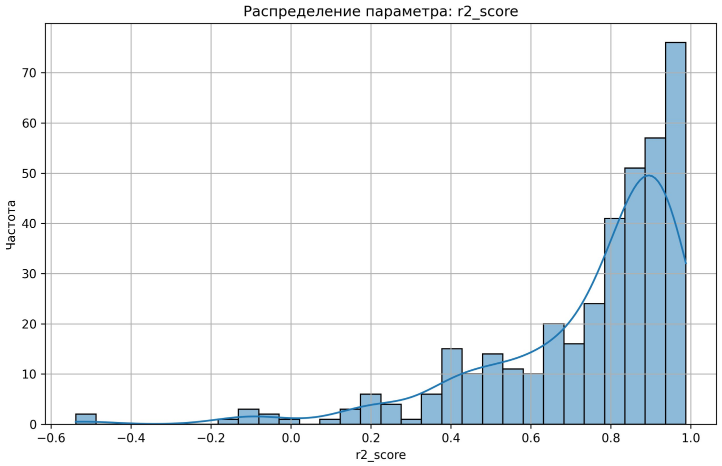


Wasserstein distance

Pearson correlation



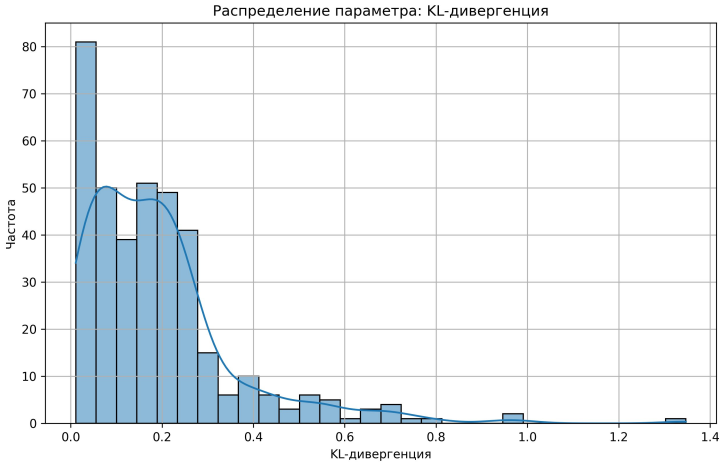
Cosine similarity



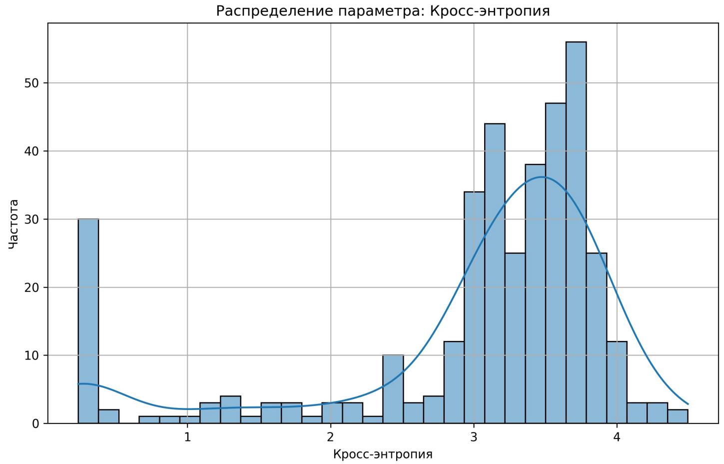
R^2

Coefficient of determination

LightAutoML results. Spectra comparison for 375 real cases

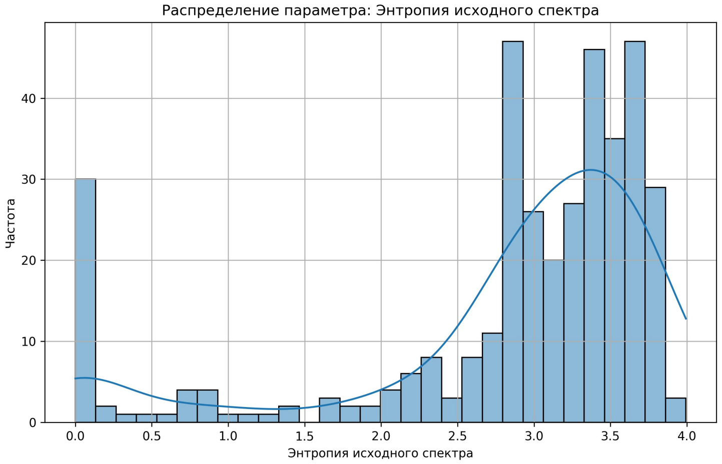


Kullback–Leibler divergence



Cross-entropy

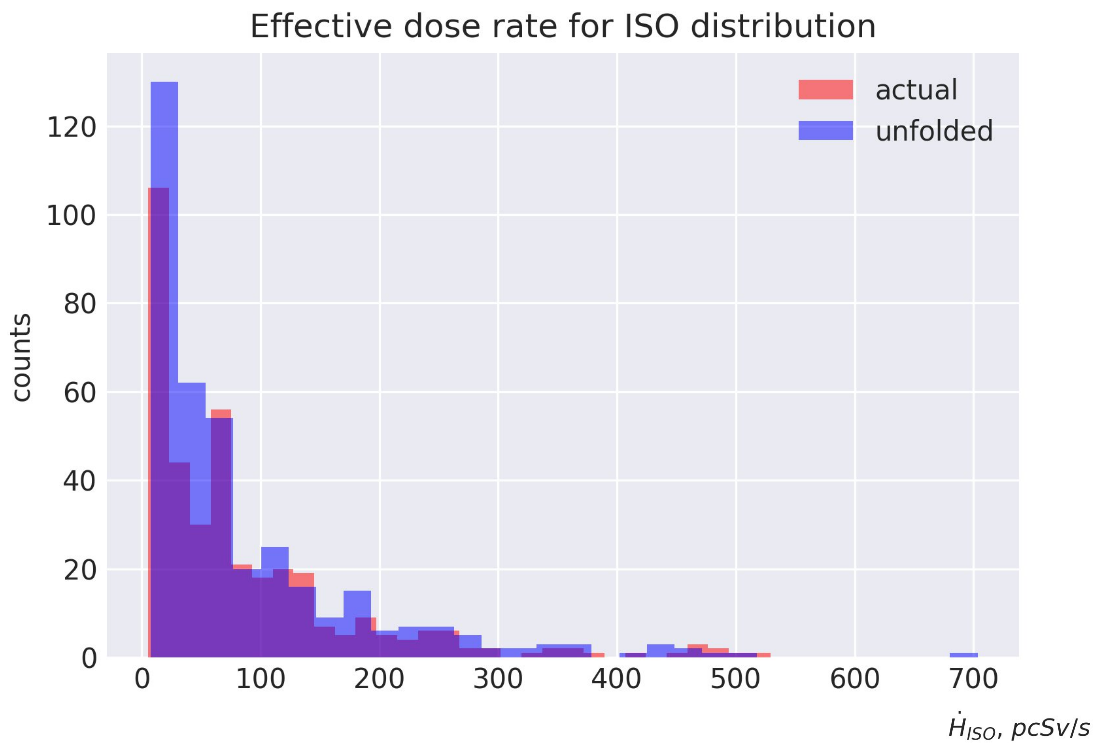
Spearman correlation

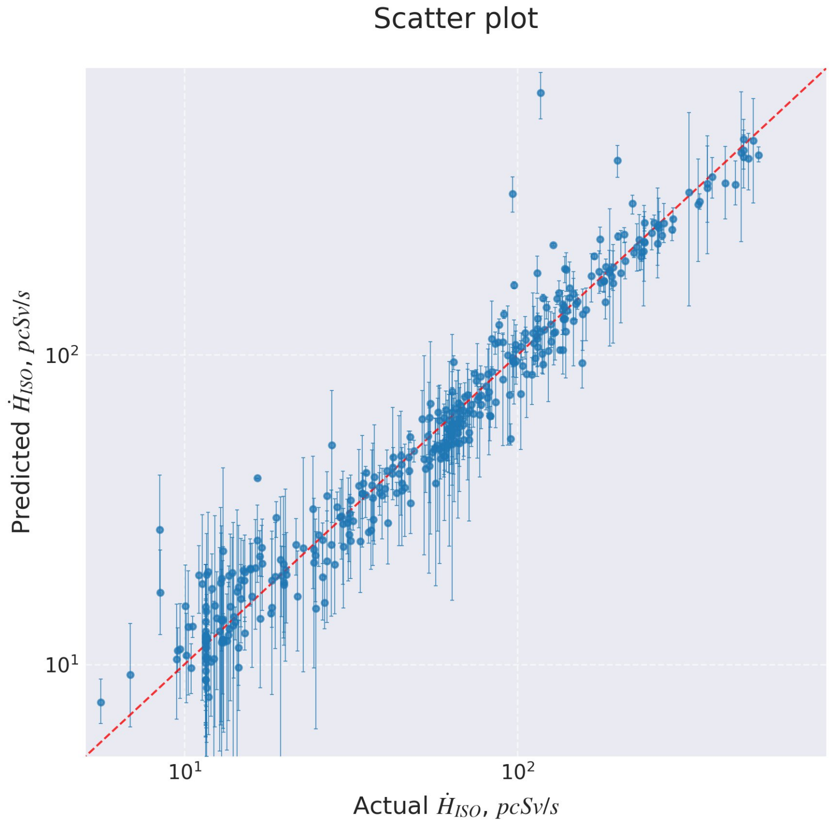


Entropy of the actual spectra

LightAutoML results. Dose assessment for 375 real spectra

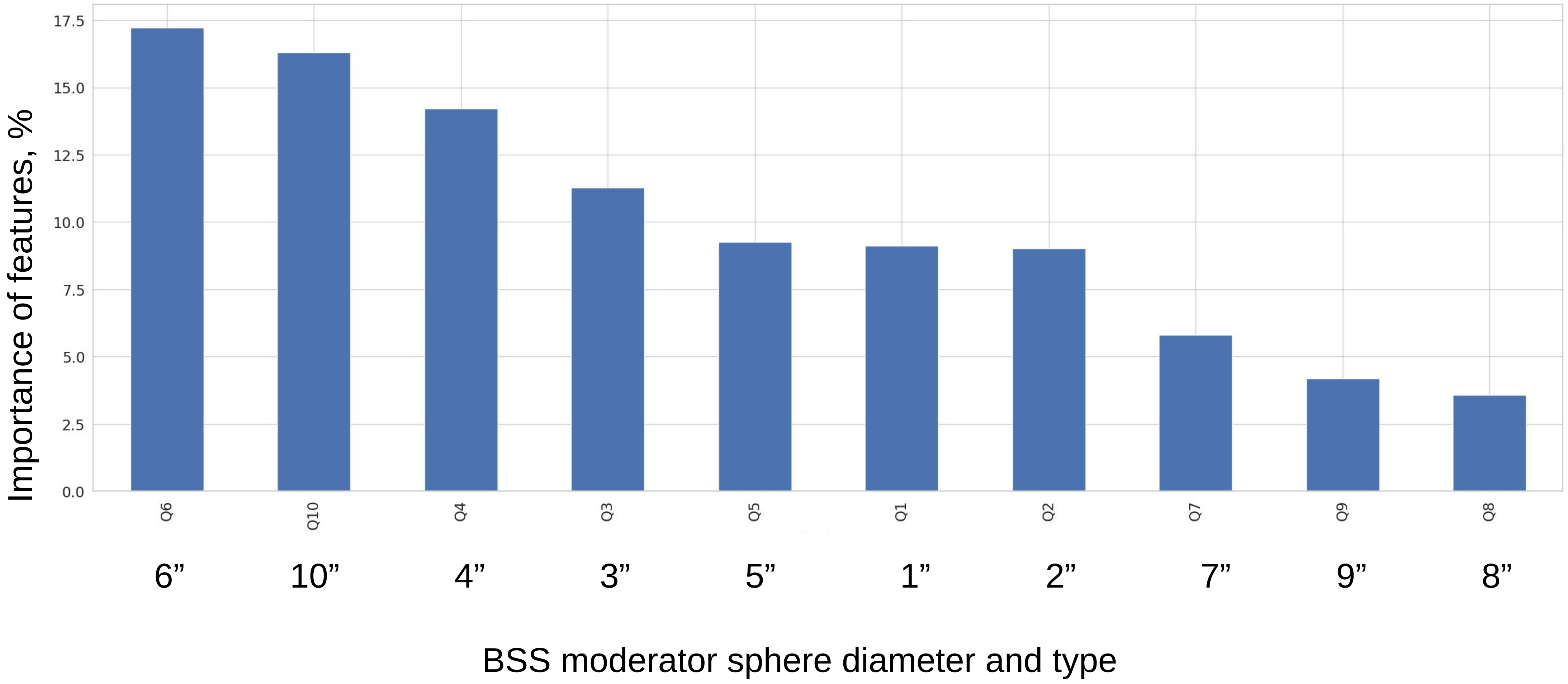
The best model selected: *Random forest regressor.*





The uncertainty of the spectra unfolding was estimated using the Monte Carlo method, in which random perturbations were introduced into the input data (measurements error = 5%).

LightAutoML. Importance of features



The method allows to select the best set of spheres for a given experimental setup based on a typical spectra.
Optimization of measurements, minimization of personnel doses.

FEDOT AutoML

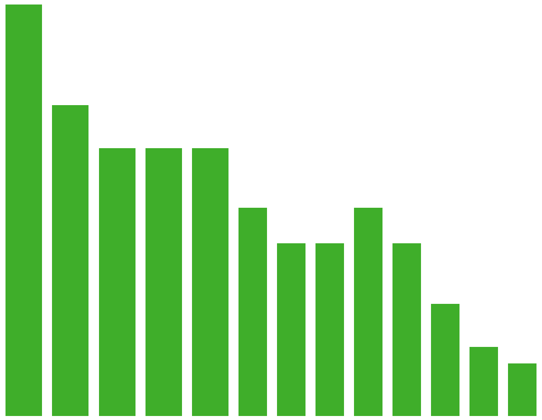
FEDOT is an open-source framework for automated modeling and machine learning (AutoML) problems.
It provides automatic generative design of machine learning pipelines for various real-world problems. The core of FEDOT is based on an evolutionary approach and supports classification (binary and multiclass), regression, clustering, and time series prediction problems.

Measurements

FEDOT

Composite model

'adareg', 'knnreg', 'lasso', 'linear', 'normalization', 'pca', 'ransac_lin_reg', 'rfr', 'ridge', 'scaling',...



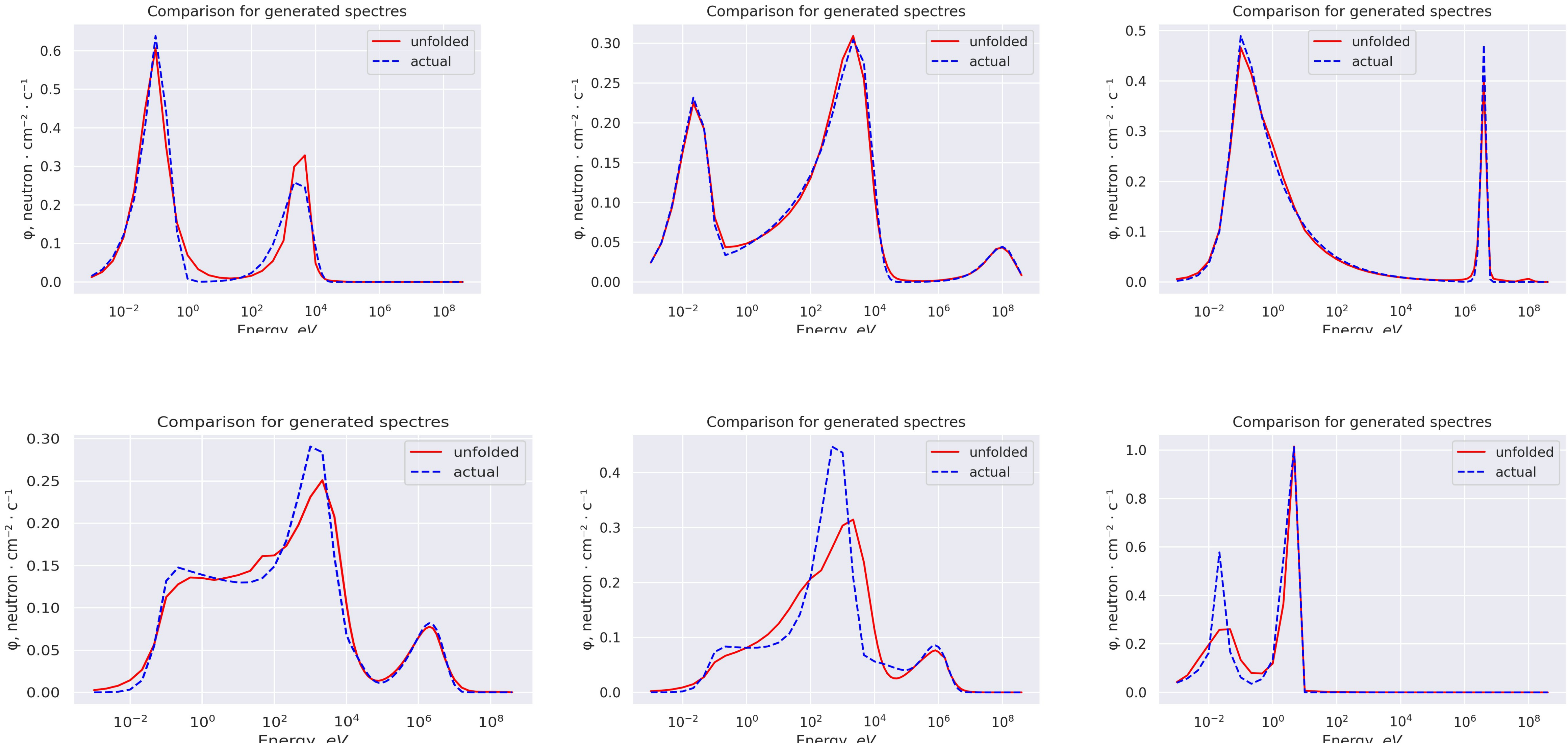
$E_1 \ldots E_S$

https://doi.org/10.1016/j.future.2021.08.022

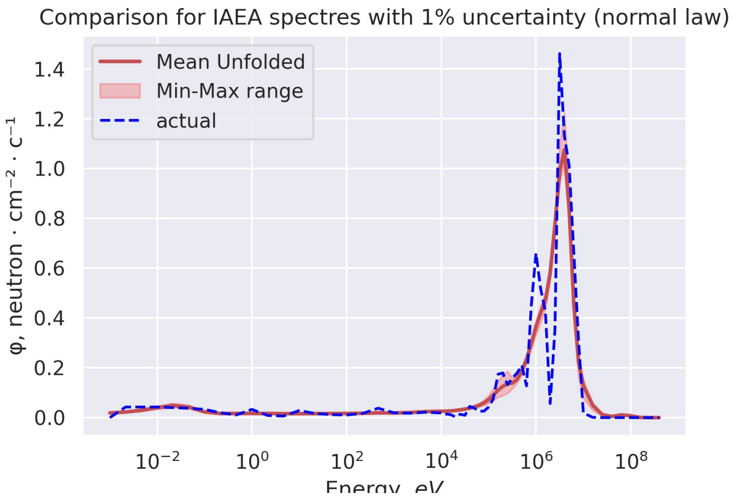
FEDOT AutoML

API name	Definition	Problem
adareg	AdaBoost regressor	Regression
catboostreg	Catboost regressor	Regression
dtreg	Decision Tree regressor	Regression
gbr	Gradient Boosting regressor	Regression
knnreg	K-nearest neighbors regressor	Regression
lasso	Lasso Linear regressor	Regression
lgbmreg	Light Gradient Boosting Machine regressor	Regression
linear	Linear Regression regressor	Regression
rfr	Random Forest regressor	Regression
ridge	Ridge Linear regressor	Regression
sgdr	Stochastic Gradient Descent regressor	Regression
svr	Linear Support Vector regressor	Regression
treg	Extra Trees regressor	Regression
xgboostreg	Extreme Gradient Boosting regressor	Regression

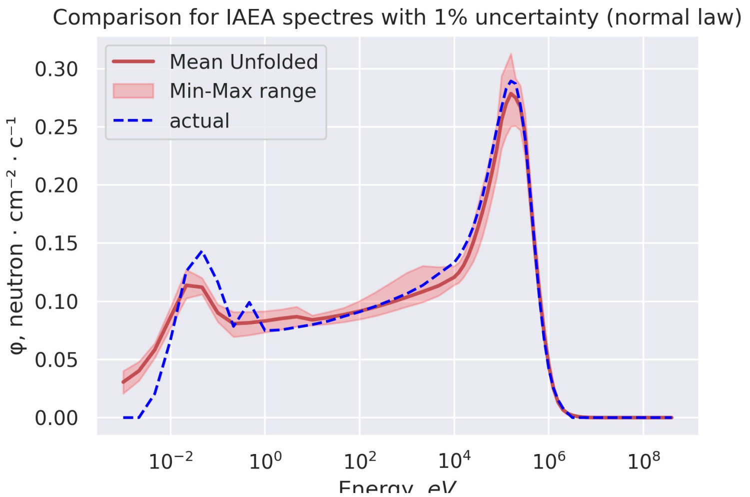
FEDOT results. Comparison for generated spectra



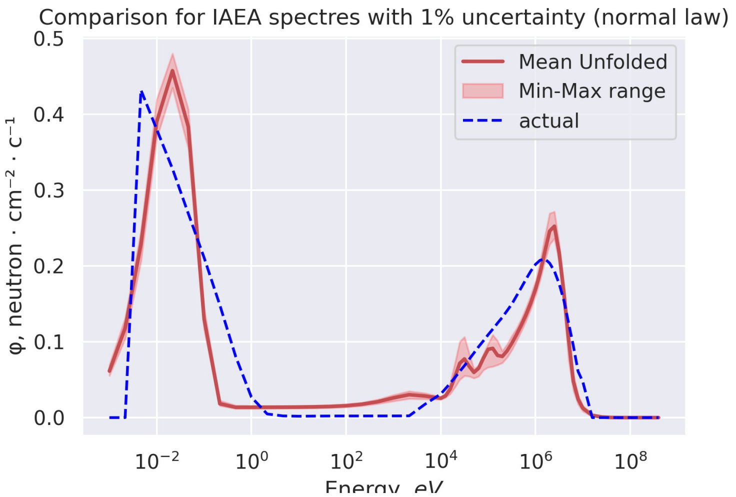
FEDOT results. Spectra comparison for 375 real cases



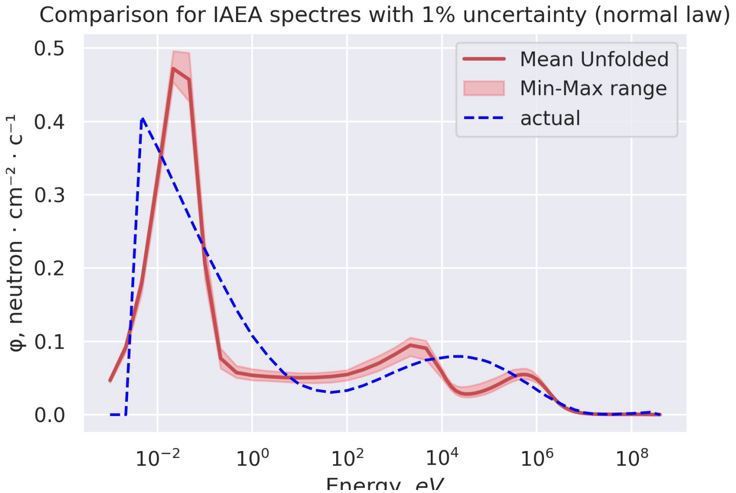
Spectra №0



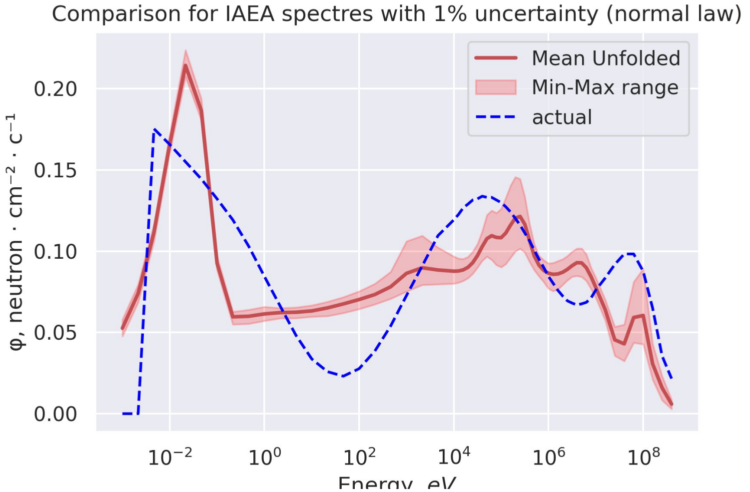
Spectra №102



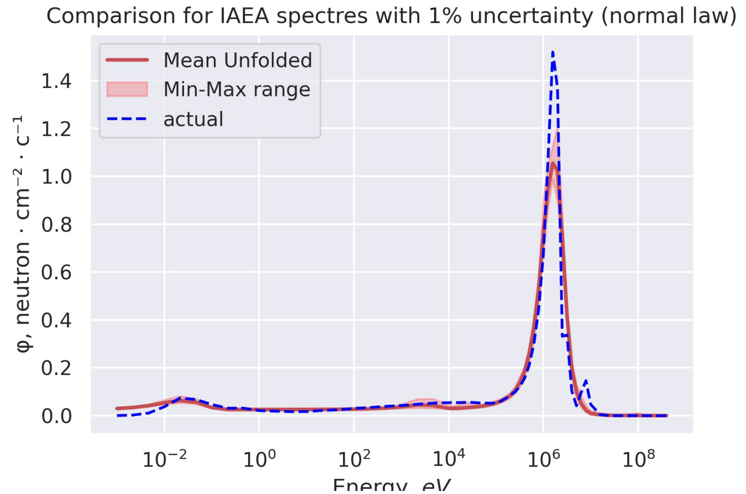
Spectra №83



Spectra №67 (soft field)



Spectra №68 (hard field)

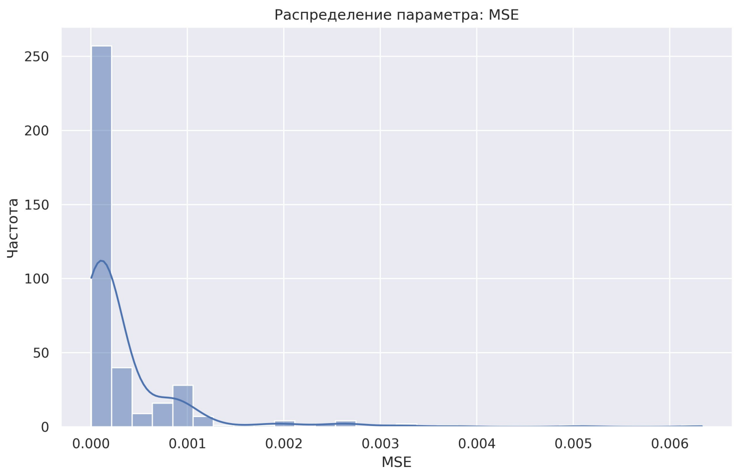


Spectra №226

The uncertainty of the spectra unfolding was estimated using the Monte Carlo method, in which random perturbations (1%) were introduced into the input data.

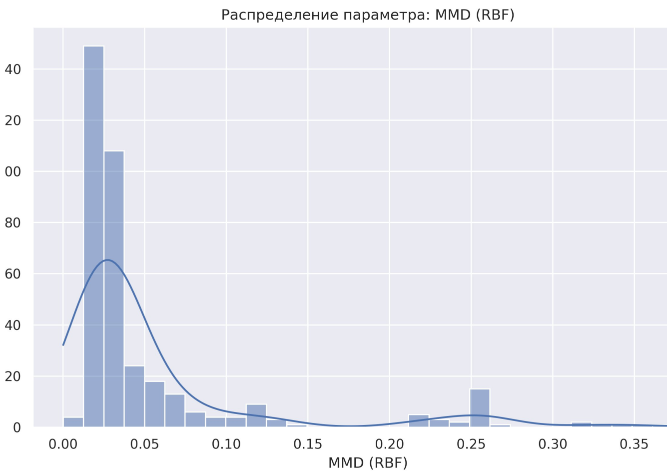
Timeout = 720 minutes

FEDOT results. Spectra comparison for 375 real cases



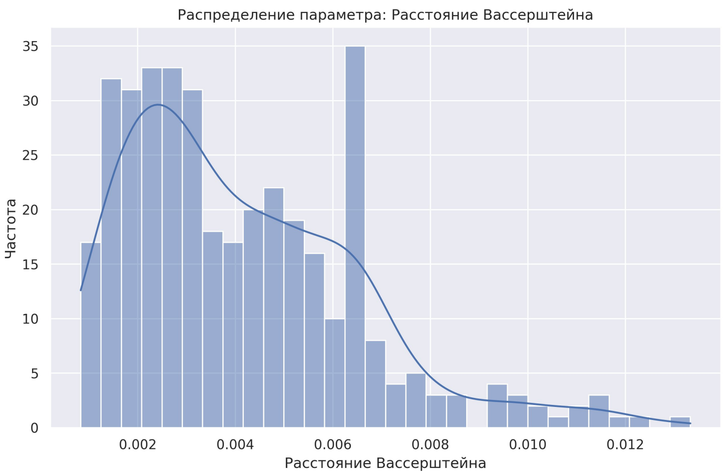
MSE

Mean squared error

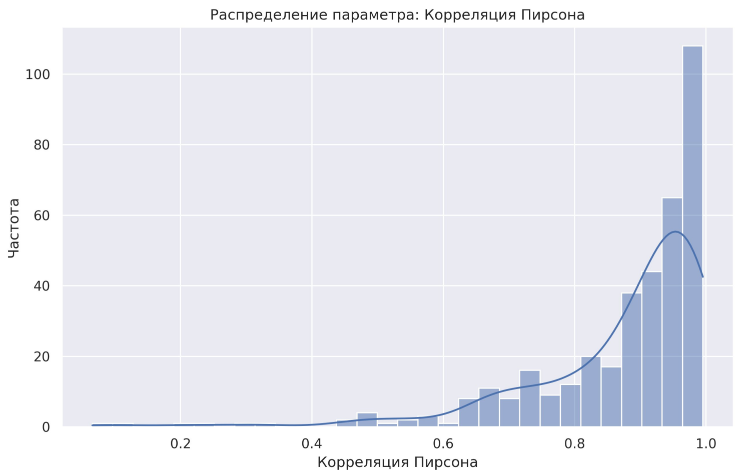


MMD

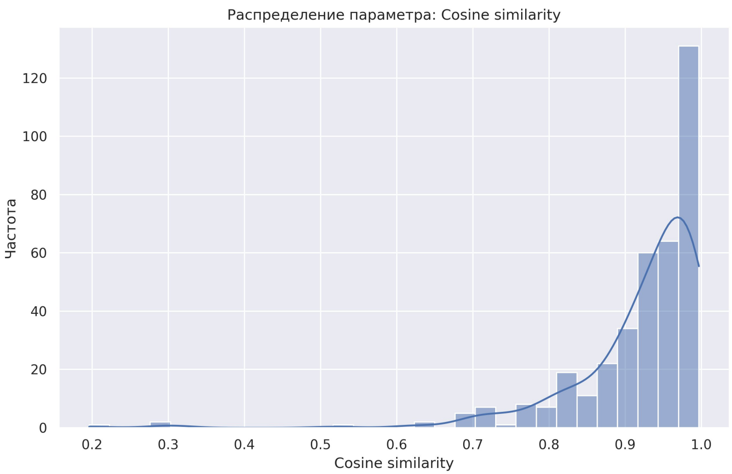
Maximum mean discrepancy



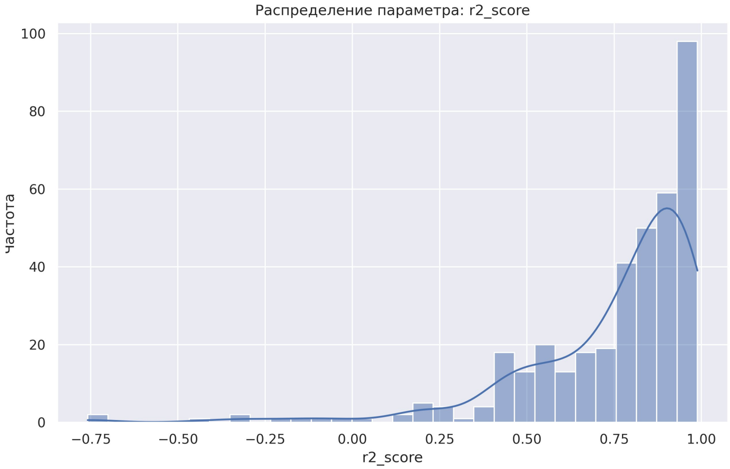
Wasserstein distance



Pearson correlation



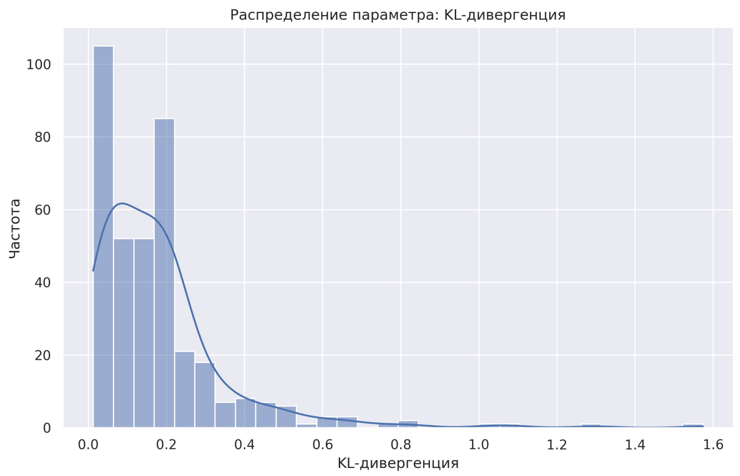
Cosine similarity



R^2

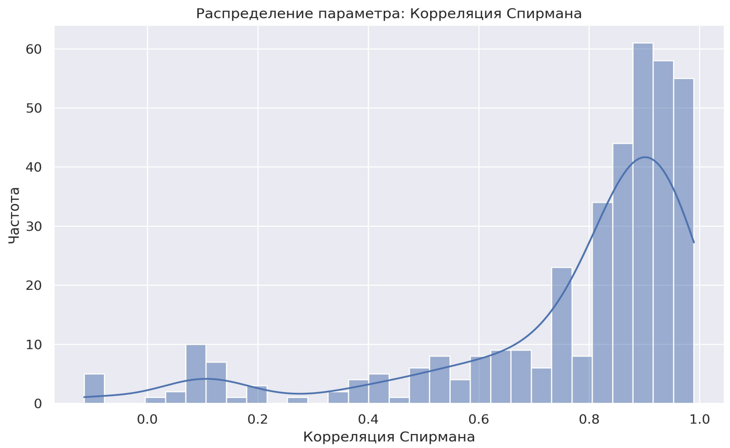
Coefficient of determination

FEDOT results. Spectra comparison for 375 real cases

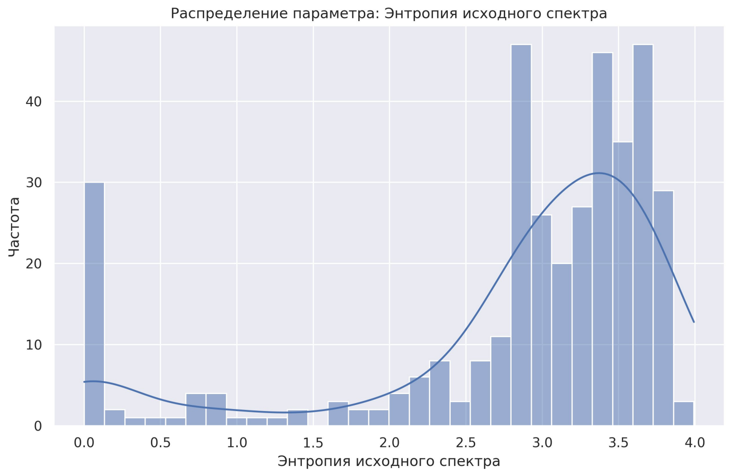


Kullback–Leibler divergence

Cross-entropy



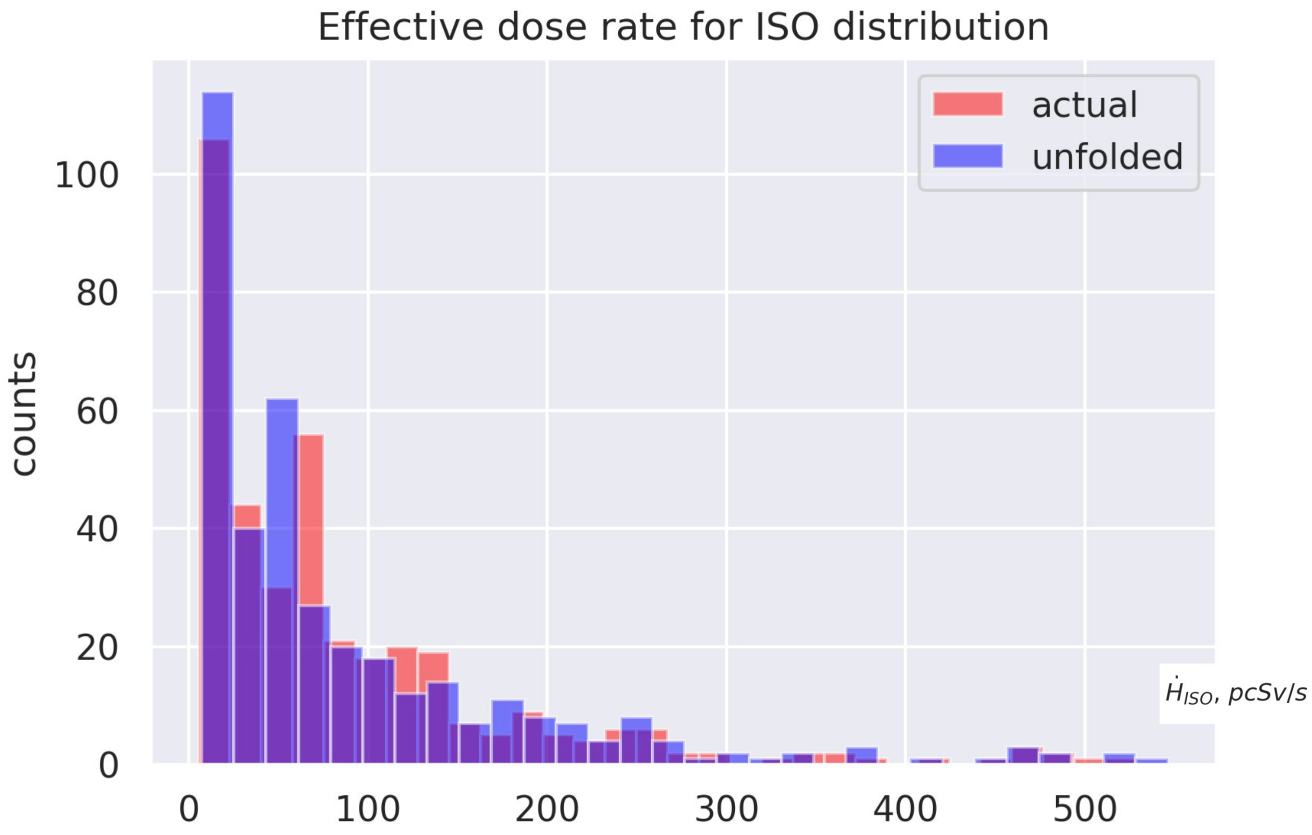
Spearman correlation

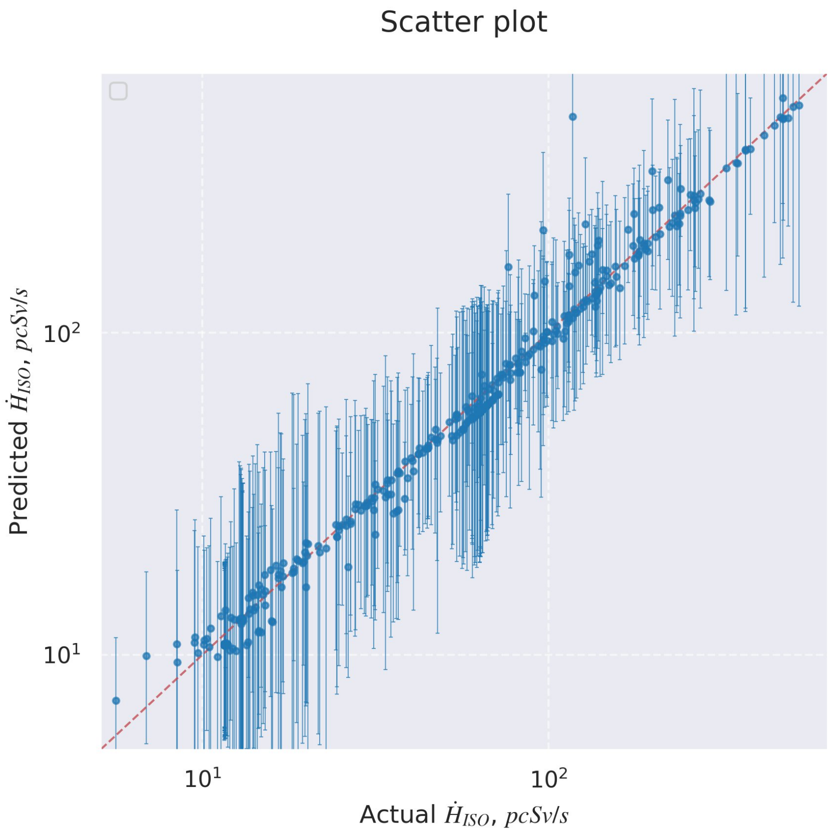


Entropy of the actual spectra

FEDOT results. Dose assessment for 375 real spectra

Timeout = 720 minutes, the best model: *Scaling + Random forest regressor.*
Timeout = 3000 minutes, the best model: *Scaling + Light Gradient Boosting Machine + Ridge Linear regressor*





The uncertainty of the spectra unfolding was estimated using the Monte Carlo method, in which random perturbations were introduced into the input data (measurements error = 1%).

Comparison of models

Mean metrics for the test dataset.

Model name	R^2	MSE	MAE	Cosine similarity	Spearman correlation	Pearson correlation
LightAutoML	0,74	3,95E-04	6,76E-03	0,916	0,744	0,866
FEDOT	0,75	3,55E-04	6,34E-03	0,920	0,772	0,873

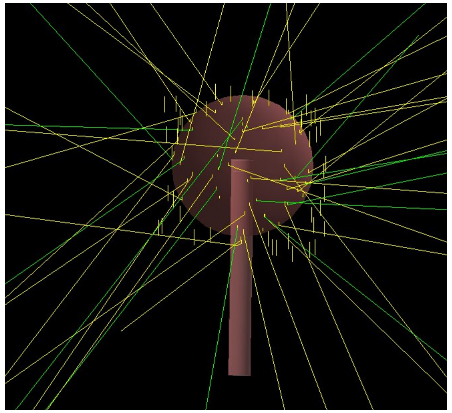
Model name	MMD	Wasserstein distance	Entropy of the actual spectra	Cross entropy	Kullback–Leibler divergence	MAPE for dose rate, %
LightAutoML	0,087	4,51E-03	2,8292	3,021	0,192	14,00
FEDOT	0,055	4,13E-03	2,8292	3,199	0,180	9,15

On a regular basis, the capabilities of nuclear enterprises around the world are tested through international exercises and must achieve within 30% error for neutron dose measurements, https://doi.org/10.1080/00295639.2025.2458437.

Discussion

Improvements of the experiment setup:

- Development of a training set using Monte Carlo[1] simulations (Geant4)[2].
- Improving the set of moderator spheres in the **high-energy region** (composite spheres with lead and other materials).
- Models for other BSS with different response functions.
- Optimal set of spheres[3,4].



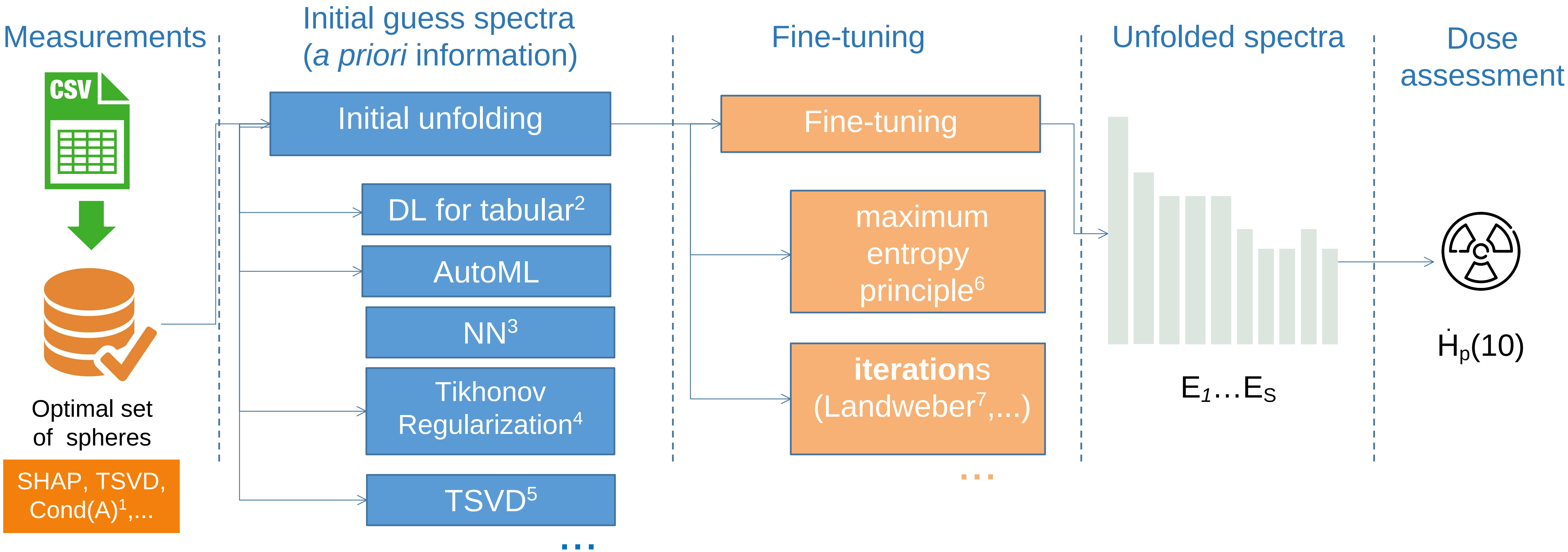
Geant4 simulation

Improvements of the spectra unfolding:

- Combination of methods to improve accuracy.
- Interpretation of the results and selecting input features with explainable artificial intelligence (XAI) methods[5].
- Other types of feature transformations[6].
- Penalty for $\varphi<0$ during training, other limitations based on the physics of neutrons.

1. Bouhadida, M. et al, Neutron Spectrum Unfolding Using Two Architectures of Convolutional Neural Networks». Nuclear Engineering and Technology 55(6),2023, 2276–82. (CNN, range 10^{-9} - 20 MeV).
2. Agostinelli S., et al. GEANT4—a simulation toolkit.*Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3,2003: 250-303.
3. Chizhov, K., Chizhov, A. Optimization of the Neutron Spectrum Unfolding Algorithm Based on Tikhonov Regularization and Shifted Legendre Polynomials. *MMCP 2024*, 74.
4. Chizhov A., Chizhov K., Unfolding of the spectra of reference neutron fields at the Phazotron (JINR) based on readings of the Bonner multi-sphere spectrometer using the truncated singular value decomposition method // LXI All-Russian Conference on Problems of Dynamics, Particle Physics, Plasma Physics and Optoelectronics, RUDN University, 2025 (in Russian).
5. Chizhov K. "Random forest regression and Shapley additive explanation for effective dose rate estimation in high-energy neutron fields based on Bonner spectrometer measurements." First Conference of Mathematics of AI, 2025, Sirius, Sochi.
6. Song W et al. Autoint: Automatic feature interaction learning via self-attentive neural networks. Proceedings of the 28th ACM international conference on information and knowledge management. 2019.

App

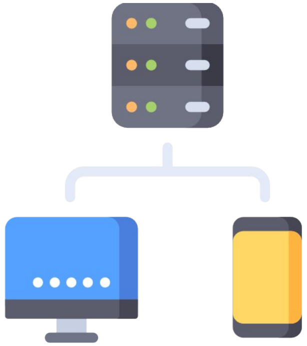


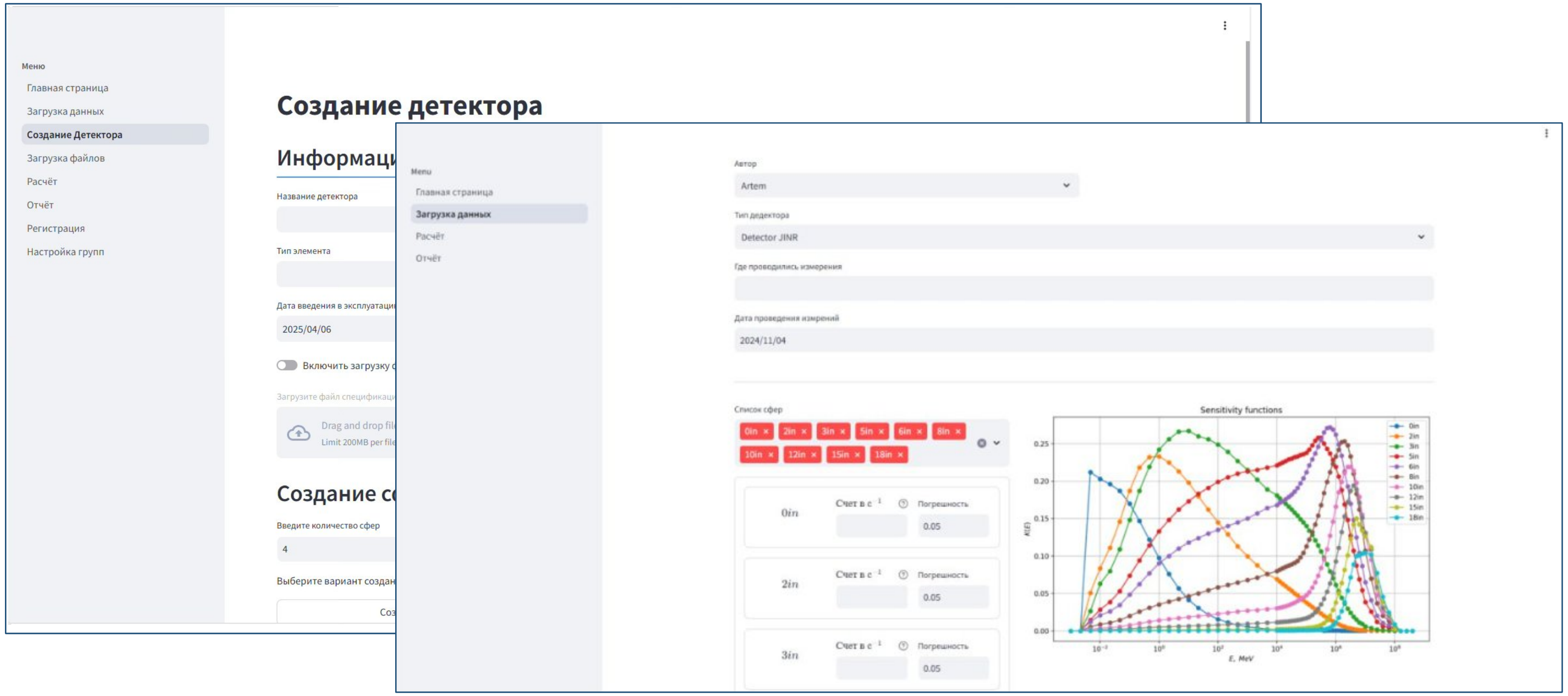
1. Chizhov K., Chizhov A. Optimization of the Neutron Spectrum Unfolding Algorithm Based on Tikhonov Regularization and Shifted Legendre Polynomials. *MMCP 2024*, 74.
2. Chizhov K., Bely A., Neutron spectrum unfolding with deep learning models for tabular data, DLCP, July, 2-4, 2025, SINP MSU, Moscow, Russia.
3. Белый А.А., Чижов К.А. Разработка веб приложения для эксперимента по восстановлению спектров нейтронов с применением алгоритма нейронных сетей, XXX Всероссийская научно-практическая конференция студентов, аспирантов и молодых специалистов, Университет "Дубна", Дубна, 2025.
4. Chizhov, K., Beskrovnaya, L., Chizhov, A. Neutron Spectra Unfolding from Bonner Spectrometer Readings by the Regularization Method Using the Legendre Polynomials, *Phys. Part. Nuclei*, 55: 532–534, 2024.
5. Chizhov A., Chizhov K., Unfolding of the spectra of reference neutron fields at the Phazotron (JINR) based on readings of the Bonner multi-sphere spectrometer using the truncated singular value decomposition method // LXI All-Russian Conference on Problems of Dynamics, Particle Physics, Plasma Physics and Optoelectronics, RUDN University, 2025 (in Russian).
6. Reginatto, M., Goldhagen P. MAXED, a computer code for maximum entropy deconvolution of multisphere neutron spectrometer data. *Health Physics* 77.5 (1999): 579-583.
7. Landweber, L.,An iteration, formula for Fredholm integral equations of the first kind. *Amer. J. Math.*73, 615–624, 1951.

Web app

A prototype of a web application for spectrum unfolding and dose assessment has been developed, features:

1. User-friendly interface.
2. User authentication.
3. Database of measurements and unfolded spectra.
4. Implementation of spectrum unfolding algorithms.
5. Cross-platform.

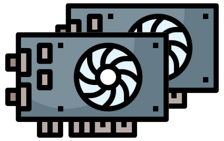




Conclusions

- This study presents implementation of automated machine learning frameworks for neutron spectrum unfolding and effective dose rate estimation using Bonner spectrometer measurements. AutoML frameworks used: LightAutoML and FEDOT.
- The application of the algorithm for the spectrum discretized into 60 points on the energy axis and expanded into 100 Legendre polynomials is demonstrated.

1. The metrics of R^2, MSE, cosine similarity, Wasserstein distance, Pearson and Spearman correlation coefficients, MMD, Kullback–Leibler divergence were assessed. For all the metrics listed, except cross-entropy, FEDOT showed a better result than LightAutoML. The difference between the metrics of the two methods is about 2%, so both methods can be used to unfold the spectrum.
2. The effective dose rate estimated from unfolded spectra showed good agreement with the actual ones, the estimation error does not exceed 14%.
3. It is proposed to use the importance of features to select the optimal set of spheres.
4. A prototype of a web application for spectrum unfolding and dose assessment has been developed.
5. The proposed method could be used for improving radiation protection in high-energy neutron fields.
6. Proposed approach could be applied to other tasks with a similar formulation of the problem.



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. 124112200072-2).

Co-authors: Bely A.A. and Starikovskaya M.D. from University “Dubna”, Dubna, Russia.

Thank you!

Publications:

1. Чижов К.А. и др. Восстановление энергетического спектра потока нейтронного излучения с помощью алгоритма машинного обучения «случайный лес». Современные информационные технологии и ИТ-образование, v. 20, n. 4, dec. 2024. ISSN 2411-1473.
2. Chizhov, K., Beskrovnaya, L. & Chizhov, A. Neutron Spectrum Unfolding Method Based on Shifted Legendre Polynomials, Its Application to the IREN Facility. *Phys. Part. Nuclei Lett.* 22, 337–340 (2025). https://doi.org/10.1134/S154747712470239X
3. 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.
4. 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.

kchizhov@jinr.ru