https://doi.org/10.61092/iaea.pd7h-zvmy

INDC(NDS)-0823
Distr. LP,G

IAEA
International Atomic Energy Agency

INDC International Nuclear Data Committee

MACHINE LEARNING FOR NUCLEAR DATA

Summary Report of the IAEA Consultants' Meeting

IAEA Headquarters, Vienna, Austria
8 – 11 December 2020
(virtual event)

D. Neudecker
Los Alamos National Laboratory (LANL)
Los Alamos, NM, USA

N. Dwivedi
Indian Institute of Technology (IIT)
Mumbai, India

E. Alhassan
Belgian Nuclear Research Centre (SCK-CEN)
Mol, Belgium

G. Schnabel
International Atomic Energy Agency (IAEA)
Vienna, Austria

October 2021

IAEA Nuclear Data Section
Vienna International Centre, P.O. Box 100, 1400 Vienna, Austria

Selected INDC documents may be downloaded in electronic form
from *http://nds.iaea.org/publications*
or sent as an e-mail attachment.
Requests for hardcopy or e-mail transmittal should be directed to
NDS.Contact-Point@iaea.org
or to:
Nuclear Data Section
International Atomic Energy Agency
Vienna International Centre
PO Box 100
1400 Vienna
Austria

Printed by the IAEA in Austria

October 2021

INDC(NDS)-0820
Distr. LP,G

MACHINE LEARNING FOR NUCLEAR DATA

Summary Report of the IAEA Consultants' Meeting

IAEA Headquarters, Vienna, Austria
8 – 11 December 2020
(virtual event)

D. Neudecker
Los Alamos National Laboratory (LANL)
Los Alamos, NM, USA

N. Dwivedi
Indian Institute of Technology (IIT)
Mumbai, India

E. Alhassan
Belgian Nuclear Research Centre (SCK-CEN)
Mol, Belgium

G. Schnabel
International Atomic Energy Agency (IAEA)
Vienna, Austria

ABSTRACT

In view of the fast advancing adoption of machine learning (ML) in various domains, a Consultants' Meeting was held to bring together researchers from nuclear physics and machine learning to present their on-going work and achievements and discuss directions of development. The presentations touched upon a broad range of topics, such as the prediction of parameters in nuclear models and the detection and localization of anomalies in nuclear power plants as well as the use of ML methodology to improve nuclear data evaluations. The wide interest in the meeting clearly indicated the promising potential of the application of ML in the field of nuclear data. To continue this unified effort, a cross-cutting IAEA Technical Meeting on Artificial Intelligence for Nuclear Technology and Applications will take place in October 2021, involving several departments of the IAEA.

October 2021

Contents

1. Introduction

The production of high-quality nuclear data libraries is a challenging undertaking. Besides human expertise, it relies on nuclear physics codes as well as evaluation and processing codes to evaluate the data of nuclear experiments and validate the resulting nuclear data files. Due to the complexities involved, it is often difficult to detect problems and trace their causes. Computational tools and procedures that assist humans in the evaluation process are therefore valuable.

Significant progress has been made in the field of machine learning (ML) over the last decade, especially regarding natural language processing and image analysis thanks to the advancement of deep learning. Along with this progress, methods of machine learning have been widely adopted in various applications and areas both to help humans make better informed decisions and to automate procedures that had to be carried out manually before.

In view of these advancements, it is pertinent to investigate how machine learning and other methods beyond standard practice, such as numerical and statistical algorithms, can help to make the nuclear data evaluation process more efficient, transparent and less error prone.

To this end, this consultants' meeting on machine learning for nuclear data was held, to review the work already done in the context of nuclear physics, to identify commonalities in the approaches, and, in general, to gain a broader perspective on potential future developments of machine learning in the field of nuclear data.

To get an overview of the status quo, both nuclear physics researchers who have already applied ML in their field of work and ML researchers were invited. Seventeen technical presentations were given, the majority of which were dedicated to the application of ML in nuclear data evaluation. The remainder mainly dealt with the improvement of machine learning methods. This report contains the presentation abstracts and, based on the work presented, a Summary and Conclusions section with a broader perspective.

Due to the COVID-19 pandemic and the concomitant travel restrictions, the meeting was held virtually. To accommodate participants in different time zones around the globe, the duration of the meeting was limited to three hours per day, between 2pm and 5pm CET. The meeting spanned four days. The first three days were dedicated to presentations and the last day was reserved for general discussions.

2. Presentation Summaries

2.1. Introduction to Consultancy Meeting on machine learning for nuclear data, G. Schnabel (IAEA)

As an introduction to the consultancy meeting, this presentation briefly reviewed the increase of compute power due to semiconductor device fabrication. Moore's law postulating that the number of transistors on microchips doubles every two years still holds today. This exponential increase in combination with the availability of large amounts of image data led to a breakthrough in image classification tasks with deep neural networks and a rapidly growing interest into the field of AI and machine learning. A definition of machine learning as the study of computer algorithms that improve automatically through experience was given. It was also emphasized that for the purpose of the meeting, the term machine learning encompasses all approaches and algorithms beyond standard practice in nuclear data and is not restricted to neural networks. The basic questions to be addressed in the context of this meeting were:

1. How can ML help to improve the handling of nuclear data?
2. How can ML help to improve the handling of nuclear models?
3. Can we learn something about physics by using ML methods?

2.2. Bayesian calibration with TMCMC, A. Gray (Liverpool Univ., UK)

Transitional Markov Monte Carlo (TMCMC) has become a popular numerical method for Bayesian calibration in uncertainty quantification since its proposal in 2007. It was originally proposed to overcome some of the issues with Metropolis-Hastings, and has the following features:

- Can sample multi-modal and high dimensional posteriors
- Importance sampling: samples are weighted and resampled
- Proposal distribution changed adaptively
- Multi-chain: each sample has an individual chain, and is simple to parallelise
- Provides Evidence: required for model class selection and model averaging

In this presentation, we discuss the theoretical details of TMCMC, with some examples. We show the use of the sampler in the 2020 Nasa UQ challenge on optimisation under uncertainty, which has a calibration problem similar to those faced in Nuclear Data evaluation (high dimensional, limited data and functional output); and how the method may be used to calibrate for sets of distributions.

2.3. When marginals or dependencies are unknown: computing with imprecise probabilities, A. Gray (Liverpool., UK)

According to Marakov [1], it was Kolmogorov who originally asked about the sum of two distributions without knowing their joint. This was answered by Marakov, who showed that the result was a set of distributions (Ref. [1]). Sklar and Frank generalised this result to other binary operations (Ref. [2]), and in this pursuit, they created copulas, a now essential object in probabilistic dependence modelling. In a seminal dissertation (Ref. [3]), Williamson described an algorithm for efficiently performing these arithmetic operations, which provide guaranteed bounds on probability distributions in terms of upper and lower cdfs. He called his method Probabilistic Arithmetic, and his sets Dependency Bounds. Since then, the method has been generalised to most of the base binary and unary operations used in programming languages. Probability Boxes (p-boxes) are the name now given to these structures, and Probability Bounds Analysis the name of the method.
P-boxes are one of many was to describe a set of probabilities, others include: intervals, possibility distributions, Dempster-Shafter structures and random sets. These structures were discovered independently, but are often synonymous and can be translated from one to another. Imprecise Probabilities links all these theories into one.
In this presentation, we review the main elements of Interval Arithmetic and Probability Bounds Analysis, showing how intervals and p-boxes bound sets of distributions, and how arithmetic operations are performed with these structures, with various degrees of knowledge about the correlations between variables. We also show that p-boxes arise naturally in situations where data is limited, or when distributional information is only partially known.

References:

[1] G.D. Makarov. Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed, Theor. Probab. Appl. **26** (1981) 803– 806.
[2] M.J. Frank, R.B. Nelsen, and B. Schweizer, Best-possible bounds for the distribution of a sum—a problem of Kolmogorov, Probab. Theory Related Fields, 74(2) (1987) 199–211.
[3] R.Ch. Williamson, Probabilistic arithmetic, PhD thesis, University of Queensland, Australia, 1989.

2.4. Validating nuclear data augmented by random forests, D. Neudecker (LANL, USA)

This talk shows how the random forest algorithm helped in exploring which nuclear data could potentially lead to differences between simulated values and experimental data of validation

measurements. This information could highlight what nuclear data should be revisited and corrected and could identify the need for future integral and differential experiments.
This question lends itself to be tackled with machine-learning (ML) methods because nuclear-data validation is a complex and highly underdetermined problem: For instance, one simulates the criticality value of one critical assembly with thousands of nuclear data from different observables of different isotopes. It is very hard to identify which specific nuclear-data value is leading to the difference between simulated values and experimental data. Traditional validation methods approach this by first fixing data of major actinides with bare spheres. Then, additional complexity is added by studying materials one at a time, such as via experiments containing reflectors of distinct materials. Due to this staggered approach of exploring specific isotopes, one is likely to overlook issues in data not studied explicitly.
We proposed in D. Neudecker et al., Nucl. Data Sheets 167, 36-60 (2020) to address this issue with random forest algorithms and the SHAP metric. It is explored whether these algorithms give expected and stable results by fabricating a sizable shortcoming in various Pu fission-source-term nuclear data and forward-propagating them to simulated criticality values. Random forest and SHAP correctly identify data of the fission-source term as culprit for bias in simulating criticality. However, they cannot distinguish between total spectrum, total fission neutron multiplicity and fission cross section as these three go together in simulating the fission-source term and cannot be disentangled - a typical issue in the input data well known to validators. In a second step it is shown that the algorithms are able to identify issues in nuclear data that would have eluded validators with traditional methods at the examples of ^{19}F(n,inl) cross sections and ^{241}Pu(n,f) cross sections in ENDF/B-VIII.0. These examples demonstrate that the random forest and SHAP algorithms not only give plausible results but also provide additional information in comparison to traditional validation methods. However, these algorithms only assist nuclear-data experts in their validation work because additional knowledge from theory and differential experimental data are needed to disentangle correlation effects due to how criticality is formulated. The ML tools shown here are already being used for every-day validation work at LANL and will contribute to validation work of the International Nuclear Data Evaluation Network (INDEN) in the future.

2.5. Highlighting physics reasons for discrepancies in differential experimental data via elastic net and random forest, D. Neudecker (LANL, USA)

This talk was about trying to understand the physics reasons leading to outliers and systematic discrepancies in an experimental database via the machine-learning (ML) algorithms elastic net and random forest. This information is valuable for the field of nuclear-data evaluation as it allows us to better quantify missing uncertainties or justify why specific data sets should be rejected.

Uncertainty quantification and outlier rejection are a fundamental part of nuclear-data evaluations including experiments. One complication in understanding what part of a measurement is driving the discrepant behavior (i.e., what part of the analysis or equipment could lead to bias) is that most nuclear-data measurements are composite in nature, i.e., the reported values (cs, nu-bar, PFNS, etc.) are extracted by combining several measurements (count rate, number of atoms in the sample, flux, etc.), and corrections. They have different attributes (features in ML language, e.g.: detector type) related to these measurements and corrections. To find the reason for outliers with expert knowledge can be tricky if there are too many common features across many experiments, but ML methods are made to find trends in a large amount of data.

To address this issue, a two-step approach is demonstrated in B. Whewell et al., NIMA Vol. 978, 164305 (2020): First outliers in the experimental data are identified with a modified version of the Hybrid Robust Support Vector Machine. In a second step, two very different machine learning methods (logistic regression with elastic net regularization and random forest regression with SHAP

feature importance metric) are used to highlight measurement features that are common among many of the outlying data points. We applied this process to features and data of 24 ^{239}Pu(n,f) cross-section data sets of the Neutron Data Standards database. We studied 37 feature categories per measurement with about 100 values across the 24 measurements. For instance, methods to determine the background, sample backing material, or impurities in the sample, are explicitly taken into account in the ML process. These input data allow us to judge whether the chosen algorithms give stable and plausible answers.

The Hybrid Robust Support Vector Machine identified data as outliers that an expert would also visually accept as such. The results of elastic net and random forest agree with regards to the most important features related to outliers and accepted data showing that the algorithms are stable for two rather different types of algorithms. It is also shown that the results are not only stable but also plausible. More specifically, the elastic net and random forest algorithms find features related to outliers and accepted data that are indeed expected by experts. But they also identify unexpected ones, showing us again that ML might highlight issues that experts can miss due to the large amount of data. If the information gained on potential physics biases in the data is translated into additional uncertainties, these added measurement uncertainties can lead to significant changes in evaluated mean values and uncertainties (within +/- 0.5% for the mean value below 10 MeV and up to 17% increase of evaluated uncertainties).

In short: Elastic net and random forest algorithms can be applied in a stable manner (i.e., enough data) and reliably (giving plausible results) to highlight potential shortcomings in the experimental data that could lead to a physics understanding of discrepancies. It is always up to the physicist to decide if the results are helpful. ML augments but does not replace expert judgment. These algorithms might be applied in the future to Neutron Data Standards data to help understand unrecognized sources of uncertainties. One issue is that we need the features for the data which is time-intensive. A collaboration with SG-50 might help.

2.6. Using machine learning to explore, diagnose, and correct for bias in nuclear data, M. Grosskopf (LANL, USA)

Misfit between integral benchmark experiments and simulation is commonly used for testing the validity of nuclear data and for directing the exploration of ways to improve nuclear data libraries. Because of the large number of isotopes and reactions relevant to these benchmarks, the space of possible relationships between this prediction bias and nuclear data sensitivities is huge. In this work we leverage the capabilities of machine learning (ML) to augment the expert-knowledge-driven search for relationships that may have been missed. To do this we build an ML predictor of bias using random forest regression to encode systematic relationships between prediction bias and nuclear data sensitivity. The ML interpretability metric SHAP is then used to quantify what sensitivities are strongly leveraged by the random forest to inform experts of potential unseen relationships driving inconsistency between simulation and experiment. This augmentation of expert knowledge can be used to identify candidate nuclear data for future adjustment using the integral experiments.

Beyond this search for relationships between nuclear data sensitivity and misfit in integral benchmark experiments, ML can be leveraged for quantifying other relationships in high-dimensional spaces for improving nuclear data. Bias in differential and integral nuclear experiments can arise from sources other than simply poorly estimated nuclear data. Experimental metadata – information about shared experimental configurations, diagnostics, etc. – can lead to outlying observations or correlated error across experiments that, if understood and accounted for, can reduce estimation bias in nuclear data. We present results using ML interpretability to quantify and communicate relationships between this

metadata and both outlying observations in differential experiments and to do nuclear data adjustment to integral experiments accounting for a sparse set of linear relationships between metadata and simulation bias.

2.7. Anomaly detection in nuclear reactors using deep learning, F. Caliva (California Univ., USA)

Monitoring nuclear reactors while running at nominal conditions is critical. Based on analysis of the core reactor neutron flux, it is possible to derive useful information for building fault/anomaly detection systems. We presented a pioneer work in the application of deep learning to detect anomalies in nuclear reactor signals. Specifically, a novel deep learning approach to unfold nuclear power reactor signals was proposed. It includes a combination of convolutional neural networks (CNN), denoising autoencoders (DAE) and k-means clustering of representations. By leveraging signal and image pre-processing techniques, the high and low energy spectra of the signals were transformed into a compatible format for CNN training. Firstly, a CNN was employed to unfold the signal into either twelve or forty-eight perturbation location sources, followed by a k-means clustering and k-Nearest Neighbour coarse-to-fine procedure, which significantly increases the unfolding resolution. Secondly, a DAE was utilized to denoise and reconstruct power reactor signals at varying levels of noise and/or corruption. The reconstructed signals were evaluated with regard to their original counter parts, by way of normalized cross correlation and unfolding metrics. The results illustrate that the origin of perturbations can be localized with high accuracy, despite limited training data and obscured/noisy signals, across various levels of granularity.

The research conducted was made possible through funding from the Euratom research and training programme 2014-2018 under grant agreement No 754316 for the 'CORe Monitoring Techniques And EXperimental Validation And Demonstration (CORTEX)' Horizon 2020 project, 2017-2021. For further details, please refer to our paper: Caliva, Francesco, De Sousa Ribeiro, Fabio, et al., A deep learning approach to anomaly detection in nuclear reactors, 2018 International joint conference on neural networks (IJCNN).

2.8. Gaussian processes under inequality constraints, F. Bachoc (Toulouse Univ., France)

Gaussian processes provide a Bayesian prior over functions. In nuclear engineering, they are widely used to address complex computer models that can be seen as black box functions. This talk will introduce Gaussian processes, with a focus on Gaussian conditioning that enables to predict the computer model values and provide confidence intervals. Then, this talk will show how constraints can be enforced in a Gaussian process model. These constraints correspond to expertise on the computer model and can be for instance boundedness, monotonicity and convexity. The talk will present the framework based on finite-dimensional approximation that guarantees to satisfy the constraints everywhere on the input space. The computational treatment of the model will be discussed, and results on simulated and nuclear test cases will be given.

2.9. Machine learning in radiation methodology: Neutron spectrum unfolding and Gamma-ray spectrometry, T. Kin (Kyushu Univ., Japan)

Machine learning approaches have been applied to radiation metrology, especially in image reconstruction of diagnosis using nuclear medicine. This trend has been observed in other fields, and

we have also adopted this approach since 2016. Currently, the following applications are investigated in our research projects:

1. The neutron spectrum unfolding ML model has been adopted especially for the multiple foils activation method. The performance is at the same level as of the conventional method, GRAVEL, but ML requires no initial guess and does not require the specification of convergence parameters.
2. Gamma-ray spectrometry: To decrease measurement time from the conventional analysis, the Covell method, we have been developing a ML model which can treat the whole gamma-ray spectrum. We found that our ML model can reduce the measurement time by 30% in a specific activity range. The range of applicability depends on the loss function used to train the ML model. We have found a new loss function for uniform performance over a wide activity range.
3. Terrestrial muon measurement: We have developed the Full-Absorption Muon Energy Spectrometer plus (FAMES+) to measure terrestrial cosmic-ray muon spectra. Two ML models will be used to eliminate other charged particles of cosmic rays, and energy prediction, respectively.

2.10. Towards the inclusion of model uncertainties in nuclear data evaluations, E. Alhassan (SCK-CEN, Belgium)

In a typical Bayesian Monte Carlo approach used in nuclear data evaluation in the fast energy region [1], parameters of pre-selected models are fine-tuned to selected experimental data. The assumption here is that the uncertainties in the nuclear data are due to our imperfect knowledge of the model parameters with the model itself being 'true' [2, 3]. Hence, by varying the parameters of these models within pre-determined uncertainties, the random cross section curves produced would overlap most of the experimental data available. This approach, however, ignores the model uncertainties and, most often, leads to a difficulty in reproducing experimental data through parameter fine-tuning alone. For example, we show in Fig. 1 that through parameter variation only not all the experimental data were covered by the random cross section curves. However, by varying many models together with their parameters, we covered all the available experimental data for the ^{58}Ni(p,2p) cross section.

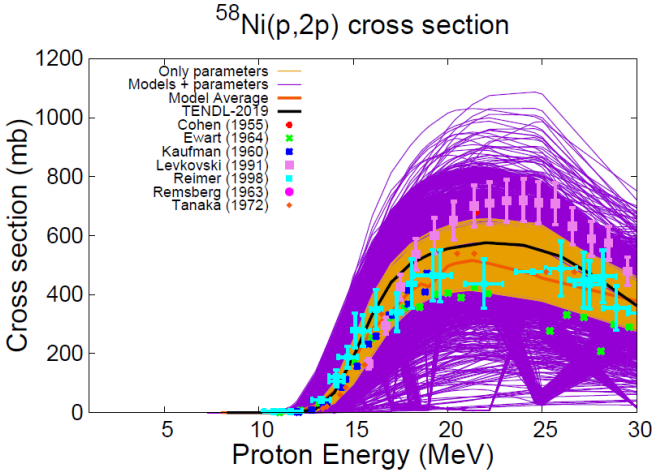

FIG. 1. Excitation functions of the ^{58}Ni(p,2p) cross section showing random cross section curves produced from the variation of many models with their parameters (purple) and only model parameters (orange), compared with experimental data from the EXFOR database. A total of 3,000 random cross section curves were produced in the case of parameter variation only, while over 10,000 were produced with the variation of many models and their parameters.

We therefore propose in this work that, instead of using a single model combination for the entire energy range of interest, we take an average over all the models within a Bayesian Model Averaging framework, using the likelihood function values as weights. Where experimental data are not available, the evaluation simply becomes the average of all the considered models. Smooth functions are then applied to the evaluation in order to smoothen-out the cross section curves. This method has been applied to p + ^{58}Ni evaluation in the fast energy region and the results obtained compared favorably with differential experimental data.

References:

[1] A.J. Koning, Bayesian Monte Carlo method for nuclear data evaluation, Eur. Phys. J. A 51 (2015) 1-16.

[2] E. Alhassan, D. Rochman, A. Vasiliev, M. Wohlmuther, M. Hursin, A.J. Koning, and H. Ferroukhi, Iterative Bayesian Monte Carlo for nuclear data evaluation, arXiv:2003.10827v1.

[3] S. Hilaire, E. Bauge, P.C. Huu-Tai and M. Dupuis, S. Péru, O. Roig, P. Romain, and S. Goriely, Potential sources of uncertainties in nuclear reaction modeling. EPJ Nuclear Sci. Technol, 4 (2018) 16.

Acknowledgement: The author would like to thank D. Rochman, A.J. Koning, G. Schnabel, H. Sjöstrand, A. Vasiliev, G. Van den Branden for useful suggestions.

2.11. Gaussian processes for treatment of model defects in nuclear data evaluations, H. Sjöstrand (Uppsala Univ., Sweden)

Work done at Uppsala University, by H. Sjöstrand, G. Schnabel, P. Helgesson, and J. Hansson was presented.

A model defect is when, independent of the choice of model parameters, a model cannot represent the actual underlying physical reality. Especially, a model with a model defect typically produces uncertainties on the model output that does not represent the magnitude of the model error. Hence, it has long been recognized that model defects can deteriorate Nuclear Data (ND) evaluations (Refs [1, 2]). Some of the features of Gaussian Processes (GPs) (Ref. [3]) were presented and how these can be used to compensate for model defects in ND evaluations. Three different options for using GPs to treat model defects were discussed. The three methods are as follows:

(1) A correction term is added to the model. This correction term is modelled with a GP. It was shown that the resulting evaluation reproduced data well. Concerns were raised on how to integrate this method in a full ND-evaluation where model constraints, e.g., sum-rules, need to be included. Regarding this method, concerns were also raised in connection with those cases where the experimental covariance is not well represented in the evaluation procedure.

(2) The GP is used to increase the flexibility of the model, f, by allowing the model parameters, $\boldsymbol{\beta}$, to become energy dependent: $f(E;\boldsymbol{\beta}+\boldsymbol{\delta}(E))$, where $\boldsymbol{\delta}(E)$ is modeled as a GP (Refs [4, 5]). Here, it was shown that the resulting evaluation reproduced the data well. Compared to method (1), this method automatically takes care of any model constraints such as sum-rules. This method gave rather large updates on the model parameters, meaning that in many cases the posterior model parameters were not consistent with the prior assumption.

(3) This method combines method (2) with adding a correction term onto the model (similar as in (1)). However, in contrast to what is done in (1), only the prior covariance of the correction term is added to the model's prior, i.e., the mean vector of the GP is ignored. The resulting evaluation also reproduced data well. In contrast to (2), this method retains more of the prior

information of the model. A hypothesis was put forward that this method might also be less sensitive to errors in the experimental covariance matrix.

Summing up, three different options to handle model defects in nuclear data evaluations using GP were presented, and arguments to use the third method were given.

References:

[1] P. Helgesson, et al., Assessment of Novel Techniques for Nuclear Data Evaluation, In *Reactor Dosimetry: 16th International Symposium*, 2018 (Eds: M. Sparks, K. DePriest, and D. Vehar) 105-116.

[2] G. Schnabel, Large Scale Bayesian Nuclear Data Evaluation with Consistent Model Defects, Ph.D thesis, TUW, 2015.

[3] C.E. Rasmussen and C.K.I. Williams, Gaussian processes for machine learning. Cambridge, Mass: MIT Press, 2006.

[4] P. Helgesson and H. Sjöstrand, Treating model defects by fitting smoothly varying model parameters: Energy dependence in nuclear data evaluation, Ann. Nucl. Energy **120** (Oct. 2018) 35–47.

[5] G. Schnabel, et al., Conception and software implementation of a nuclear data evaluation pipeline, Nucl. Data Sheets **173** (2021) 239-284, http://arxiv.org/abs/2009.00521.

2.12. AI/ML-based evaluation in the resonance region, V. Sobes (Tennessee Univ., USA)

The work currently on-going at the University of Tennessee on the application of Artificial Intelligence algorithms to nuclear data evaluation in the resonance region was presented. This work will be realized in the form of a machine-learning-based wrapper for the SAMMY resonance evaluation code. The advantages of an AI/ML-based approach to nuclear data evaluation were highlighted such as introducing reproducibility into nuclear data evaluation, greatly accelerating the time necessary to produce a new evaluation and providing reliability and systematic uncertainty quantification in the process. The approach currently pursued by Sobes and his collaborators is limited to automating the evaluation process after all of the experimental data has been compiled by a human evaluator. One of the challenges faced by the AI/ML-algorithm approach is the fact that statistical models of the experimental uncertainty are hardly ever reported and always assumed to be Gaussian. The group is currently researching methods for the algorithm to be self-aware of its own limitations and to identify regimes where even the estimate of uncertainty may not be reliable. Sobes concluded with the observation that even with rigorous statistical models for uncertainty quantification, AI/ML-based evaluation of nuclear data may raise concerns about the interpretability of results and decisions made by the algorithm.

2.13. GPyTorch: Gaussian processes for modern machine learning systems, G. Pleiss (Columbia Univ., USA)

Gaussian processes (GPs) are powerful machine learning models that offer well-calibrated uncertainty estimates, interpretable predictions, and the ability to encode prior knowledge. They are commonly used in many applications ranging from geo-statistics to health care to robotics. At the same time, GPs suffer from practical limitations: poor scalability, challenging implementations, and inefficient use of modern compute hardware.

The GPyTorch library is a Gaussian process framework specifically designed to combat these practical issues. Taking inspiration from neural networks, the GPyTorch library construct inference procedures that only use matrix multiplication and other operations amenable to GPU acceleration. This not only

makes inference more scalable and efficient, but it also simplifies the implementation of specialty models. It will be demonstrated that GPyTorch scales exact GP models to datasets with millions of observations – two orders of magnitude larger than existing frameworks. Many common GP models can be implemented in less than 30 lines of Python code.

2.14. Nuclear data generation using Gaussian process regression and its related topics, H. Iwamoto (JAEA, Japan)

Gaussian processes are used as a Bayesian machine learning technique to solve regression and classification problems. Because a regression model based on Gaussian processes (GPs) gives predictive distributions for the regression problems, this technique may be a powerful tool not only for evaluating nuclear data but also for determining the corresponding uncertainty.

In the previous study, a method to generate nuclear data from experimental data was proposed based on GPs, and it was demonstrated that our method can generate reasonable regression curves and their uncertainties from experimental data. In this meeting I presented this technique as well as some merits and issues for the nuclear data generation.

2.15. Uninformative, sparse, smooth: GP priors on second derivatives of cross sections, G. Schnabel (IAEA)

Gaussian process regression is a flexible non-parametric method for supervised learning. In nuclear data evaluation, Gaussian processes (GPs) have already been used (1) to fit cross sections [1], (2) as an energy-dependent prior on model parameters [2], and (3) to account for the deficiencies of a nuclear model in an evaluation, e.g., [3,4]. The time requirement to fit a Gaussian process scales with the cube of the number of data points, which makes it problematic to apply them to large data sets. At the time of writing, data sets with more than ten thousand data points may be considered large. This presentation outlined a specific Gaussian process construction which scales better to large sets. The first ingredient is to approximate the full Gaussian process by imposing a multivariate normal distribution on a finite mesh and to employ linear interpolation to obtain the full Gaussian process specification. The second ingredient is to impose the multivariate normal distribution with a diagonal covariance matrix on the second derivative of the cross section as prior. This construction introduces a high degree of sparsity into the Gaussian process, which can be exploited by relying on sparse matrix algorithms.

References:

[1] H. Iwamoto, Generation of nuclear data using Gaussian process regression, J. Nucl. Sci. Technol. 57/8 (2020) 932-938

[2] P. Helgesson, H. Sjöstrand, Treating model defects by fitting smoothly varying model parameters: Energy dependence in nuclear data evaluation, Ann. Nucl. Energy, 120 (2018) 35-47.

[3] H. Leeb, D. Neudecker, Th. Srdinko, Consistent Procedure for Nuclear Data Evaluation Based on Modeling, Nucl. Data Sheets 109/12 (Dec. 2008) 2762-2767.

[4] G. Schnabel, H. Leeb, Differential Cross Sections and the Impact of Model Defects in Nuclear Data Evaluation, EPJ Web of Conferences 111 (Mar 2016) 09001.

2.16. Trees, forests and islands: A machine learning approach to nuclear physics, N. Dwivedi (IIT, India)

Machine Learning (ML) and Artificial Intelligence (AI) have found a place in our contemporary technology to discover patterns, and to classify and predict any large data. These classifications and

predictions find applications from social media, online behaviour to banking, stock market movement, etc. Recent use of ML in high energy physics for search of exotic particles, and making faster computations in molecular dynamics calculations, etc., has opened the vistas of their usage in fundamental sciences.

In this work we explore such a ML algorithm for model generation and prediction of various nuclear observables like damping parameters, shell correction energies, quadrupole deformations, pairing gaps, level densities and giant dipole resonances for a large number of nuclei. We, in particular, predict level density parameters for superheavy elements which is of great current interest.

We train Gradient Boosted Trees and Random forests on the available nuclear data to obtain a prediction model. We use as features the energy terms of the semi empirical mass formula derived from the liquid drop model. The performance of the prediction model is good. This work is detailed in https://arxiv.org/abs/1907.09764

2.17. Machine learning for neutron resonance evaluations, D. Brown (BNL, USA)

The performance of nuclear reactors and other nuclear systems depends on a precise understanding of the neutron interaction cross sections for materials used in these systems. These cross sections exhibit a resonance structure whose shape is determined in part by the angular momentum quantum numbers of the resonances. The correct assignment of the quantum numbers of neutron resonances is therefore of paramount importance. In this presentation, we describe the application of machine learning to automate the quantum number assignments. Scikit-learn classifiers were trained on simulated resonance data whose statistical properties were chosen to mimic real data. We explored the use of several physics (and random matrix theory)-motivated features for training the classifiers, including the nearest neighbor spacing distribution, cumulative level distribution, and channel width distributions. Initial results demonstrated that we can determine resonance spin groups somewhat reliably. We are now investigating the application of our approach to ^{52}Cr resonance data.

2.18. Probabilistic machine learning for uncertainty quantification, A. Lovell (LANL, USA)

The use of machine learning (ML) is becoming more prominent in the nuclear theory community. However, their use for uncertainty quantification (UQ) usually relies on ML algorithms implemented as emulators with a separate UQ formalism built on top. To directly extract uncertainties from machine learning, we instead explore a probabilistic ML algorithm, the Mixture Density Network (MDN). Through the MDN, instead of using a standard neural network to directly map from input to output, the output is described as a sum of Gaussians and the weight, mean, and variance of each Gaussian is determined through the neural network. In this way, uncertainties from the training set – taken from the nuclear data observables of interest – can be propagated to the predicted observables. Here, we explore the use of the MDN to reproduce fission mass yields and their uncertainties. Using spontaneous fission of ^{252}Cf, we study how many training samples are needed for converged results, how different levels of uncertainties propagate from the input to output, and how well the normalization and symmetry of the yields are upheld. In addition, using energy-dependent mass yields of neutron-induced fission of ^{235}U, we explore the ability of the MDN to interpolate and extrapolate between and beyond incident energies in the training set. Although further studies are still needed, this initial investigation indicates that the MDN is a promising machine learning algorithm for use in predicting nuclear data with well-quantified uncertainties.

2.19. Constrained Bayesian optimization of criticality experiments at LLNL, D. Siefman (LLNL, USA)

The design of criticality experiments is typically an iterative process that employs a Monte Carlo transport code. The goal is to find a design that optimizes some variable, like the sensitivity of a response to a cross section, while simultaneously ensuring criticality. The high fidelity of the Monte Carlo code is a great asset, but it makes exploring the design space computationally expensive. Herein, we present how a constrained Bayesian optimization algorithm can be used to efficiently design a criticality experiment. It uses Gaussian processes as a surrogate model to probe the design space and to reduce the number of code executions that are needed to find the optimum. We demonstrate constrained Bayesian optimization with a U-235/polyethylene spherical system and a TEX experiment that is designed for criticality safety validation of a nuclear waste model at the Hanford Site. For both systems, a global optimum was found within 75 Monte Carlo simulations.

3. Summary and conclusions

The presentations indicated that ML approaches are already being pursued in the nuclear science and engineering domain and first applications have proven successful. Hence recent advancements in ML methodology are likely to benefit applications in the nuclear field. We conclude by summarizing and highlighting the achievements outlined in participants' presentations regarding (1) nuclear data evaluation and validation, (2) other applications in nuclear science and engineering, and (3) general ML methodology.

(1) In nuclear data evaluation, the widely adopted Generalized Least Squares (GLS) method to combine experimental data for an evaluation is sensitive to discrepant data and outliers and their proper identification is therefore important. D. Neudecker and M. Grosskopf elaborated on a two-step procedure to determine problematic experimental features related to outliers by using Support Vector Regression (SVR) in combination with either elastic nets or random forests. M. Grosskopf further highlighted the SHAP metric to extract feature importance from classifiers that are not easily interpretable, such as random forests, and presented an extension of the GLS method that also accounts for experimental biases. This approach can guide future measurements and highlight features of experiments that should be explored to maximally impact future evaluations rather than strive to reduce uncertainties of a single measurement - a costly and difficult task.

E. Alhassan and H. Sjöstrand addressed in their presentations the issue that in model-based evaluations nuclear models are not always able to follow the experimental data closely. The adoption of model-based evaluations therefore may introduce significant biases in nuclear-data libraries and addressing them is of high importance from the application point of view as often the model defect uncertainties are not reflected in the evaluated uncertainties. E. Alhassan presented a solution based on Bayesian model averaging to combine several nuclear physics models to obtain a better overlap between the ensemble of model predictions and the experimental data. H. Sjöstrand presented two options to integrate Gaussian processes into a model-based evaluation to obtain more reliable uncertainties and better agreement between the evaluation and the experimental data, which either impose GPs as priors on energy-dependent model parameters or add them to the model prediction. In a preliminary study, he also presented results using the two approaches in combination with a favorable impact on evaluated uncertainties. H. Iwamoto reviewed GP regression and elaborated on the possibility to evaluate nuclear data with GP regression instead of using a nuclear physics model and

showed several evaluations of various proton-induced reactions obtained by GP regression. This approach is potentially less prone to biases compared to model-based evaluations but requires decisive experimental data. G. Schnabel outlined in his presentation a GP construction that scales well to a large number of data points and enforces smooth solutions, which is often desirable for cross sections curves in the fast energy range. Depending on the specifics of the evaluation, several hundreds of thousands of data points are feasible. N. Dwivedi trained Gradient Boosted Trees (GBT) to predict nuclear properties, such as level density parameters, using the terms in the Bethe-Weizsäcker mass formula as features. A. E. Lovell presented results with density mixture networks to predict important fission observables, such as mass yields. A notable feature of density mixture networks is that uncertainties are also learned and predicted by the approach. Approaches that quantify uncertainties are relevant for the ongoing CRP on "Updating Fission Yield Data for Applications" as covariances for fission yields are often not provided.

In the resolved resonance range, evaluations are performed by using R-matrix fits, which require the correct assignment of several quantum numbers. The proper assignment is often achieved by trial and error and considerable human intervention to obtain good fits to experimental data. D. Brown presented work that aims to automate, among other quantities, the spin assignment by using synthetic data with known spin assignment as training data to train various ML classifiers available in the scikit-learn Python module.

Nuclear data libraries are validated and tweaked to perform favorably in predicting integral experiments. This step is important for quality assurance as it guarantees that the libraries can be reliably used for applications represented by the integral experiments. D. Neudecker explained how random forests in combination with the SHAP metric can be used to trace potential problems in nuclear data. The approach used features associated with the experiments, such as the reflector material, and the sensitivity coefficients that link the simulated experimental values to specific nuclear data to learn to predict discrepancies between simulated and experimental values of various integral responses. The SHAP metric was then used to identify problematic measurement features and nuclear data that might need to be revised after careful analysis by the evaluator.

Besides leveraging the information of existing benchmarks, new benchmarks sensitive to specific aspects of the nuclear data can be built to help constrain nuclear data or to identify problematic aspects of an evaluation. D. Siefman introduced Bayesian global optimization as a tool to find good designs of integral benchmark experiments. A key ingredient of this approach is to replace the computationally expensive transport model by a computationally inexpensive Gaussian process surrogate model that mimics the original model. This surrogate model is then be used in the design optimization. This line of research is important because previously integral experiments were designed with expert judgement mapping out the best design space; humans may miss possible and good designs.

(2) Regarding the use of ML in the broader nuclear science and engineering domain, T. Kin reported several successful applications of ML in radiation metrology. He elaborated on neutron spectrum unfolding in the multiple-foils activation method using a feed-forward neural network, and on the determination of the activity of Caesium-137 and other isotopes from a gamma-ray spectrum using convolutional neural networks. He also talked about the application of ML in the measurement of terrestrial cosmic-ray muon spectrum. F. Caliva explained how deep neural networks can be trained on data generated with a simulation code (CORE SIM) to perform anomaly detection in nuclear reactor cores. Noise analysis today, which is seldomly

applied, requires an analyst to look at the neutron noise measurements, and guess what is happening. As this unfolding is far from trivial, using AI for the unfolding represents a significant advancement.

(3) The remaining presentations focused on pure statistics and ML methodology. A. Gray presented the Transitional Markov Chain Monte Carlo (TMCMC) approach to efficiently sample from posterior distributions with potentially multiple peaks. This approach may have the potential to improve uncertainty quantification in nuclear data evaluation. He also outlined an alternative paradigm to Bayesian statistics to compute with probabilities, which are probability boxes (or p-boxes in short). G. Pleiss addressed the issue that standard algorithms for Gaussian process regression cannot work with a large number of data points. He outlined improvements in numerical methods to enable Gaussian process inference also in situations with large amounts of data. In view of the recent uses of Gaussian processes in nuclear data evaluation, this methodological improvement indicates that global evaluations with GPs across the nuclide chart may become possible soon. F. Bachoc showed that inequality constraints can be incorporated into GP regression, which is also highly relevant for nuclear data evaluation. Constraints, such as cross sections being non-negative, are usually not formally taken into account in the commonly employed Generalized Least Squares method to produce an evaluation.

It was an interesting observation that neural networks were mostly used in situations where it is possible to generate training data by running a model with varied model parameters, such as in the work on anomaly detection in reactor cores presented by F. Caliva and on spectrum unfolding presented by T. Kin. The purpose of the neural networks in these scenarios was to learn the inverse function, i.e., to estimate the parameters of the model based on the predictions. In the other cases GP regression and random forests were employed.

In summary, presentations gave evidence that ML has already been applied beneficially in nuclear science and engineering and its use can be expected to expand in the future thanks to the improvements of compute hardware and advances in ML methodology.

APPENDIX I: ADOPTED AGENDA

all times in CET

San Francisco	CET-9
New York	CET-6
Japan	CET+8

Connecting via Webex will be possible from 13:50 onwards every meeting day
All presentation times comprise 5 min of Q&A at the end

Tuesday 8 December

Time	Duration	Name	Affiliation	
14:00	0:10	A. Koning	IAEA	Welcome address
14:10	0:15	G. Schnabel	IAEA	Introduction
14:25	0:05			Election of chair and rapporteur
				Presentations by Participants
14:30	0:20	A. Gray	ULIV	Bayesian calibration with Transitional Markov Chain Monte Carlo
14:50	0:20	A. Gray	ULIV	When marginals or correlations are unknown: UQ with imprecise probabilities
15:10	0:20	D. Neudecker	LANL	Validating nuclear data augmented by Random Forests
15:30	*0:10*			*Break*
15:40	0:20	D. Neudecker	LANL	Highlighting physics reasons for discrepancies in differential experimental data via Elastic Net and Random Forest
16:00	0:25	M. Grosskopf	LANL	Using Machine Learning to explore, diagnose, and correct for bias in nuclear data
16:25	0:25	F. Caliva	UCSF	Anomaly detection in nuclear reactors using deep learning
16:50	0:30			**Discussion**

Wednesday 9 December

Time	Duration	Name	Affiliation	**Presentations by Participants (cont'd)**
14:00	0:35	F. Bachoc	IMT/UPS	Gaussian processes under inequality constraints
14:35	0:25	T. Kin	Kyudai	ML in radiation metrology: Neutron spectrum unfolding and Gamma-ray spectroscopy
15:00	0:25	E. Alhassan	SCK-CEN	Towards the inclusion of model uncertainties in nuclear data evaluations
15:25	*0:10*			*Break*
15:35	0:25	H. Sjöstrand	UU	Gaussian Processes for treatment of model defects in nuclear data evaluations
16:00	0:15	V. Sobes	UTK	ML wrapper around Sammy
16:15	0:25	G. Pleiss	CU	GPyTorch: Gaussian processes for modern machine learning systems
16:40	0:30			**Discussion**

Thursday 10 December

Time	Duration	Name	Affiliation	Presentations by Participants (cont'd)
14:00	0:25	H. Iwamoto	JAEA	Nuclear data generation using Gaussian process regression and its related topics
14:25	0:25	G. Schnabel	IAEA	Uninformative, sparse and smooth: GP priors on second-derivatives of cross sections
14:50	0:25	N. Dwivedi	IIT Bombay	Trees, Forests and Islands - ML approach to nuclear physics
15:15	*0:10*			*Break*
15:25	0:25	D. Brown	BNL	Machine Learning for neutron resonance evaluations
15:50	0:25	A. Lovell	LANL	Probabilistic Machine Learning for Uncertainty Quantification
16:15	0:25	D. Siefman	LLNL	Constrained Bayesian optimization of integral experiments at LLNL
16:40	0:30			**Discussion**

Friday 11 December

Time	Duration	
14:00	3:00	**General discussion and report preparation**

APPENDIX II: PARTICIPANTS

Country	Name	Surname	Affiliation	Email
BELGIUM	Erwin	ALHASSAN	SCK-CEN	erwin.alhassan@sckcen.be
CHINA	Junchen	PEI	Peking University	peij@pku.edu.cn
FRANCE	Francois	BACHOC	University Toulouse	francois.bachoc@math.univ-toulouse.fr
	Jean-Christophe	David	CEA Saclay	jean-christophe.david@cea.fr
	Cyrille	DE SAINT JEAN	CEA	cyrille.de-saint-jean@cea.fr
INDIA	Nishchal	DWIVDI	Indian Institute of Technology Bombay	dwivedi.nishchal@gmail.com
JAPAN	Tsunenori	INAKURA	Tokyo Institute of Technology	inakura.t.aa@m.titech.ac.jp
	Chikako	ISHIZUKA	Tokyo Institute of Technology	chikako@lane.iir.titech.ac.jp
	Yoshihisa	TAHARA	Tokyo Institute of Technology	taharayo@mc.point.ne.jp
	Hiroki	IWAMOTO	Japan Atomic Energy Agency	iwamoto.hiroki@jaea.go.jp
	Tadahira	KIN	Kyushu University	kin.tadahiro.981@m.kyushu-u.ac.jp
	Ryohei	TAKAHASHI	Kyushu University	takahashi.ryouhei.759@s.kyushu-u.ac.jp
	Shoto	WATANABE	Hokkaido University	shoto@nucl.sci.hokudai.ac.jp
SWEDEN	Erik	BRANGER	Uppsala University	erik.branger@physics.uu.se
	Henrik	SJÖSTRAND	Uppsala University	henrik.sjostrand@physics.uu.se
	Andreas	SOLDERS	Uppsala University	andreas.solders@physics.uu.se
	Christophe	DEMAZIERE	Chalmers University of Technology	demaz@chalmers.se
UNITED KINGDOM	Ander	GRAY	Liverpool University	ander.gray@liverpool.ac.uk

Country	Name	Surname	Affiliation	Email
UNITED STATES	David	BROWN	Brookhaven National Laboratory	dbrown@bnl.gov
	Gustavo	NOBRE	Brookhaven National Laboratory	gnobre@bnl.gov
	Michael	GROSSKOPF	Los Alamos National Laboratory	mikegros@lanl.gov
	Amy	LOVELL	Los Alamos National Laboratory	lovell@lanl.gov
	Denise	NEUDECKER	Los Alamos National Laboratory	dneudecker@lanl.gov
	Ruby	ARAJ	Lawrence Livermore National Laboratory	araj1@llnl.gov
	Catherine	PERCHER	Lawrence Livermore National Laboratory	percher1@llnl.gov
	Daniel Jerome	SIEFMAN	Lawrence Livermore National Laboratory	siefman1@llnl.gov
	Francesco	CALIVA	University of California	francesco.caliva@ucsf.edu
	Geoff	PLEISS	Columbia University	gmp2162@columbia.edu
	Vladimir	SOBES	University of Tennessee	sobesv@utk.edu
INT. ORGANIZATION	Michal	FLEMING	OECD Nuclear Energy Agency	michael.fleming@oecd-nea.org
	Georg	SCHNABEL	International Atomic Energy Agency	g.schnabel@iaea.org
	Roberto	CAPOTE	International Atomic Energy Agency	r.capotenoy@iaea.org
	Matteo	BARBARINO	International Atomic Energy Agency	m.barbarino@iaea.org
	Kalle	HEINOLA	International Atomic Energy Agency	k.heinola@iaea.org
	Christian	HILL	International Atomic Energy Agency	ch.hill@iaea.org
	Arjan	KONING	International Atomic Energy Agency	a.koning@iaea.org
	Ludmila	MARIAN	International Atomic Energy Agency	l.marian@iaea.org
	Shin	OKUMURA	International Atomic Energy Agency	s.okumura@iaea.org
	Naohiko	OTSUKA	International Atomic Energy Agency	n.otsuka@iaea.org
	Elmira	PIRSHAYAN	International Atomic Energy Agency	e.pirshayan@iaea.org

APPENDIX III: PRESENTATION LINKS

#	Author	Title	Link
1	G. Schnabel	Introduction to Consultancy Meeting on Machine Learning for Nuclear Data	pptx
2	A. Gray	Bayesian Calibration with Transitional MCMC	pptx
3	A. Gray	When marginals or dependencies are unknown: computing with imprecise probabilities	pptx
4	D. Neudecker	Validating Nuclear Data Augmented by Random Forests	pdf
5	D. Neudecker	Highlighting Physics Reasons for Discrepancies in Differential Experimental Data via Elastic Net and Random Forest	pdf
6	M. Grosskopf	Using Machine Learning to Explore, Diagnose, and Correct for Bias in Nuclear Data	pdf
7	F. Caliva	Anomaly detection in nuclear reactors using deep learning	pdf
8	F. Bachoc	Gaussian processes under inequality constraints	pdf
9	T. Kin	Machine learning in radiation metrology: Neutron spectrum unfolding and Gamma-ray spectrometry	pptx
10	E. Alhassan	Towards the inclusion of model uncertainties in nuclear data evaluations	pdf
11	H. Sjöstrand	Gaussian Processes for treatment of model defects in nuclear data evaluations	pptx
12	V. Sobes	AI/ML-based evaluation in the resonance region	pdf
13	G. Pleiss	GPyTorch: Gaussian Processes for Modern Machine Learning Systems	pdf
14	H. Iwamoto	Nuclear data generation using Gaussian process regression and its related topics	pptx
15	G. Schnabel	Uninformative, sparse, smooth: GP priors on second derivatives of cross sections	pdf
16	N. Dwivedi	Trees, Forests and Islands: A Machine learning approach to Nuclear Physics	pptx
17	D. Brown	Machine Learning for Neutron Resonance Evaluations	pdf
18	A. Lovell	Probabilistic Machine Learning for Uncertainty Quantification	pdf
19	D. Siefman	Constrained Bayesian Optimization of Criticality Experiments at LLNL	pptx
20	G. Schnabel	Opening of final day of CM on ML for ND	pptx

Nuclear Data Section
International Atomic Energy Agency
Vienna International Centre, P.O. Box 100
A-1400 Vienna, Austria

E-mail: nds.contact-point@iaea.org
Fax: (43-1) 26007
Telephone: (43-1) 2600 21725
Web: http://nds.iaea.org