NC STATE UNIVERSITY

Artificial Intelligence and Machine Learning for Scientific Computing in Nuclear Engineering: Status, Benchmarks, and Outlook

Xu Wu
xwu27@ncsu.edu

Assistant Professor
https://www.ne.ncsu.edu/artisans/
Department of Nuclear Engineering
North Carolina State University

Webinar on "Year in Review: Nuclear Thermal Hydraulic Achievements of 2023"
Thermal Hydraulics Division, American Nuclear Society

Friday, March 15, 2024

Outline

Outline

- Recent performance breakthroughs in AI/ML, especially advances in DL, the availability of powerful, easy-to-use ML libraries (e.g., scikit-learn, TensorFlow, PyTorch.), and increasing computational power and data have led to unprecedented interest in AI/ML among nuclear engineers.
- AI/ML is attracting increasing interest in the nuclear thermal-hydraulics area:
  1. ATH-2018: panel
  2. ATH-2022: panel, 1 session with 4 papers
  3. NURETH-19 (2021): keynote, 3 technical sessions with 13 papers
  4. NURETH-20 (2023): keynote, panel, 6 technical sessions with 25 papers

  and multiple technical sessions in ANS Meetings.
- Other activities:
  1. PHYSOR-2022: 4 technical sessions with 15 papers
  2. M&C-2023: 6 technical sessions with 28 papers
  3. PHYSOR-2024: 2 technical sessions with 10 papers
  4. "SciML for NE Applications" Workshop series in M&C and PHYSOR Conferences since 2021
  5. INL AI/ML Symposium[1] Series 1.0 (04/2020) - 12.0 (11/2023)
  6. U.S. NRC "Data Science and AI Regulatory Applications" Workshops[2] #1 (06/2021) - #4 (09/2023)

  and more...

[1] https://inl.gov/artificial-intelligence/

[2] https://www.nrc.gov/public-involve/conference-symposia/data-science-ai-reg-workshops.html

A list of some representative application areas

1. data-driven closure models in nuclear thermal-hydraulics,
2. data-driven material discovery and qualification for nuclear materials,
3. data-driven autonomous operation and control for advanced reactors,
4. data-driven diagnosis and prognosis,
5. DT-based predictive maintenance, oversight and compliance monitoring of the operations of a nuclear power plant, and design of advanced reactors,
6. data-driven structural health monitoring for nuclear systems,
7. nuclear data evaluation,
8. reactor core/assembly design optimization,
9. nuclear power plant cyber-attack detection,
10. ML-based surrogate models for calibration, validation, uncertainty quantification, sensitivity analysis and optimization,
11. data-driven prediction based on high-fidelity simulations,
12. data-driven prediction based on measurement data,

and many more...

"SciML for Nuclear Engineering Applications" workshop series in M&C and PHYSOR conferences

Year	Conference	Speaker	Institution	SciML Workshop Titles
2021	M&C	Xu Wu	NCSU	Uncertainty Quantification and Scientific Machine Learning
2021	M&C	Majdi Radaideh	MIT	NeuroEvolution Optimization with Reinforcement Learning
2021	M&C	Han Bao	INL	A Machine Learning Approach for Scale Bridging in System-level Thermal-hydraulic Simulation
2021	M&C	Massimiliano Fratoni	UC Berkeley	Machine Learning Augmented Cross Section Evaluation
2021	M&C	Yang Liu	ANL	Physics-Informed Machine Learning
2022	PHYSOR	Som Dhulipala	INL	Active learning for computational simulations: Application to TRISO fuel failure analysis
2022	PHYSOR	Ayman Hawari	NCSU	Development of Neural Thermal Scattering (NeTS) Modules for Data Representation and Applications
2022	PHYSOR	Linyu Lin	INL	Development of a Nearly Autonomous Management and Control System for Advanced Reactors
2022	PHYSOR	Vladimir Sobes	UTK	Applications of AI/ML from Nuclear Data to Reactor Design
2022	PHYSOR	Justin Watson	U of Florida	Prediction of PWR Pin Powers using Convolutional Neural Networks
2023	M&C	Xu Wu	NCSU	Everything You Need to Know About Neural Networks
2023	M&C	Brendan Kochunas	Umich	Hybrid Machine Learning Techniques in Microreactor Control
2023	M&C	Stephen Lam	U Mass Lowell	ML-Enhanced Interrogation of Complex Material Structures and Properties
2023	M&C	Hany Abdel-Khalik	Purdue	Inferring with Intelligence: A Nuclear Perspective
2024	PHYSOR	William Gurecky	ORNL	Latent Neural Controlled Differential Equations for Time Series Forecasting
2024	PHYSOR	Ryan McClarren	Notre Dame	Machine learning for neutronics applications
2024	PHYSOR	Syed Bahauddin Alam	UIUC	Intelligent Digital Twin for Nuclear Systems
2024	PHYSOR	Diego Mandelli	INL	Causal inference as a bridge between simulation models and observed phenomena
2024	PHYSOR	Mohammad Abdo	INL	Sparse Sensing and Sparse Learning for Nuclear Digital Twins

Issues and gaps

- The full potential of AI/ML techniques has not been fully realized in NE. One barrier is that existing ML methods developed for **low-consequence error-tolerant settings** (e.g., image classification, recommender system) often do not meet the needs of **high-consequence nuclear applications**.
- Some gaps:
  1. Application-agnostic algorithms, or those designed for traditional ML applications such as computer vision, natural language processing, speech recognition, and audio synthesis, where the datasets are in the format of images, words, audios and videos, cannot be directly applied to scientific data in NE without non-trivial, task-specific modifications.
  2. Verification, validation and uncertainty quantification (VVUQ) of AI/ML,
  3. Data scarcity - more challenging for NE,
  4. AI/ML algorithmic transparency (explainability, interpretability),
  5. AI/ML robustness (reproducibility, applicability),
  6. AI/ML privacy/confidentiality,
  7. Regulatory acceptance of AI/ML deployment,

  and more...

Outline

1. Status of AI/ML in NE

2. OECD/NEA Task Force on AI/ML and Benchmarks

3. Outlook on Near Term Topics

Outline

Examples on how benchmarks have accelerated AI/ML developments

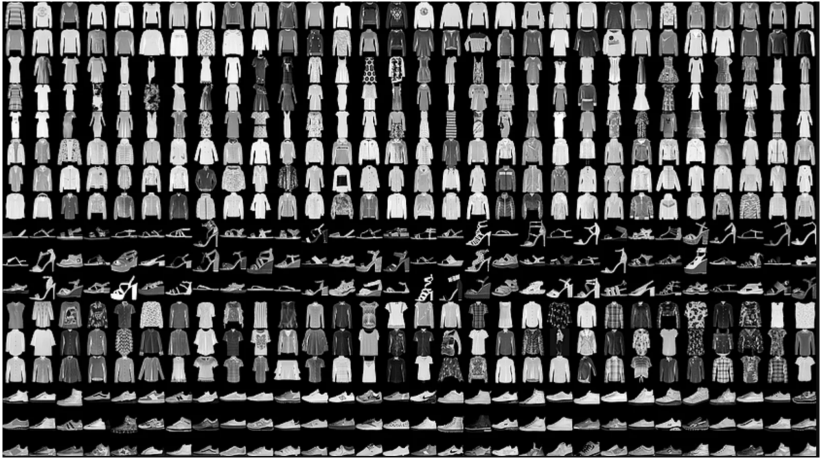
- Benchmarks have greatly promoted the success of ML/DL algorithms in the last two decades. For example: the MNIST[3] dataset for handwritten digits recognition, and Fashion MNIST[4].

[3] http://yann.lecun.com/exdb/mnist/

[4] https://www.kaggle.com/datasets/zalando-research/fashionmnist

Examples on how benchmarks have accelerated AI/ML developments

- Benchmarks have greatly promoted the success of ML/DL algorithms in the last two decades. For example: the MNIST[3] dataset for handwritten digits recognition, and Fashion MNIST[4].



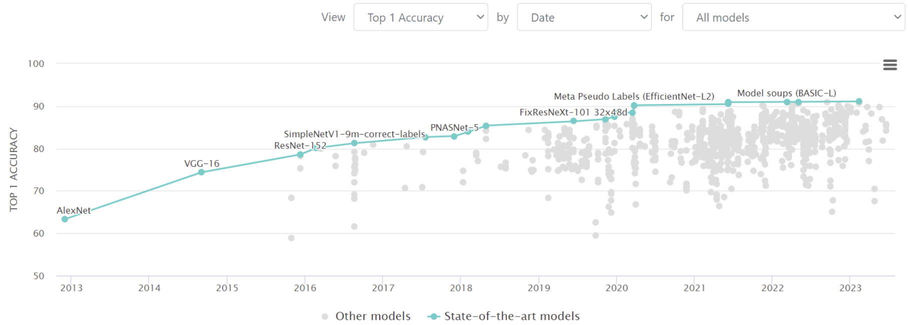
[3] http://yann.lecun.com/exdb/mnist/

[4] https://www.kaggle.com/datasets/zalando-research/fashionmnist

- Another example is ImageNet[5]. The ImageNet project is a large visual database designed for use in visual object recognition software research.
- More than **14 million images** have been hand-annotated by the project to indicate what objects are pictured. ImageNet contains more than **20,000 categories**.
- During 2010-2017, the ImageNet project runs an annual software contest, the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), where software programs compete to correctly classify and detect objects and scenes.

[5] https://www.image-net.org/

- Another example is ImageNet[5]. The ImageNet project is a large visual database designed for use in visual object recognition software research.
- More than **14 million images** have been hand-annotated by the project to indicate what objects are pictured. ImageNet contains more than **20,000 categories**.
- During 2010-2017, the ImageNet project runs an annual software contest, the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), where software programs compete to correctly classify and detect objects and scenes.



[5] https://www.image-net.org/

Outline

- The "Task Force on AI and ML for Scientific Computing in Nuclear Engineering" was initiated on June 2022 at the 15th OECD/NEA WPRS Workshops, under the EGMUP.
- This Task Force aims at developing benchmark exercises to evaluate the performance of AI/ML in multi-physics Modeling and Simulation (M&S) of reactor systems.
- The benchmark exercises will span various computational domains on which the participants could develop and evaluate the performance of their ML methods.
- The benchmark will focus on establishing and improving ML predictive credibility, through rigorous verification, validation and uncertainty quantification (VVUQ) of AI/ML which matches the quality standards for VVUQ of traditional nuclear M&S models.
- It will provide recommendations to the WPRS and the nuclear community on the scientific development needs (data, methods, and evaluation standards) for trustworthy AI/ML applications in nuclear scientific computing problems.

Outline

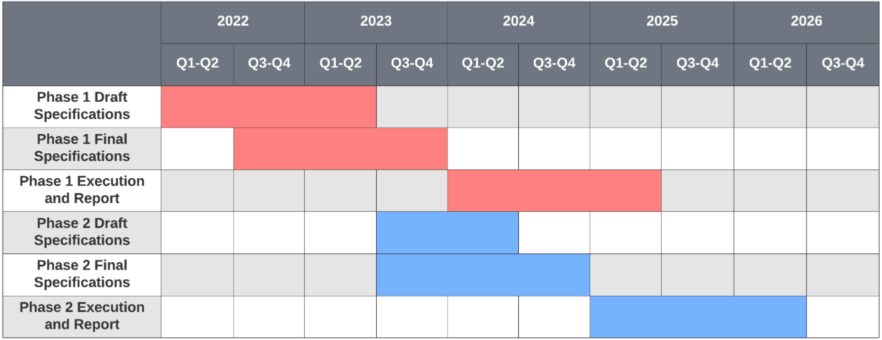
Task Force scope

- This Task Force will design and execute AI/ML benchmark exercises relevant to various computational domains (e.g. reactor physics, thermal-hydraulics, fuel performance, multi-physics).
- Certified experimental data, as well as verified and validated high-fidelity computational data from other WPRS expert groups will be leveraged as training dataset for the benchmark exercises.
- The AI/ML benchmark will be organized in exercises based on specific AI/ML tasks:
  1. Supervised ML: regression and classification,
  2. Unsupervised ML: dimensionality reduction and clustering,
  3. VVUQ of AI/ML models,
  4. Anomaly detection,
  5. Transfer learning,
  6. Deep generative AI for data augmentation,
  7. Optimization based on deep reinforcement learning.

  and more...

Benchmark exercise timeframe

- The development and execution of the benchmark exercises will be in two phases. Phase 1 will contain standard AI/ML tasks, while Phase II will target at more advanced AI/ML tasks.
- Phase 1: will mainly focus on tasks related to regression, classification, dimensionality reduction, clustering, and anomaly detection.
- Phase 2: will mainly focus on tasks related to VVUQ, transfer learning, deep generative learning and deep reinforcement learning.



Task Force deliverables

1. Develop specification of benchmark exercises, including overview of both traditional and novel AI/ML techniques, existing experimental data and high-fidelity computational data (both single physics and multi-physics) for benchmarking of AI/ML methodologies.
2. Prepare reports on the results of benchmark exercises, as well as multi-national efforts on establishing and improving the AI/ML predictive credibility through rigorous VVUQ. Conduct webinar on conclusions form comparative analysis and lessons learned based on the report.
3. Provide benchmark updates on the benchmark activities in the yearly WPRS Workshops.
4. Organize yearly education workshops, short courses, and webinars on fundamental and advanced AI/ML techniques. Convert the above-mentioned benchmark exercises to educational benchmark exercises and made them available at the OECD/NEA Gitlab repositories and demonstrate their use within the proposed education activities.

Outline

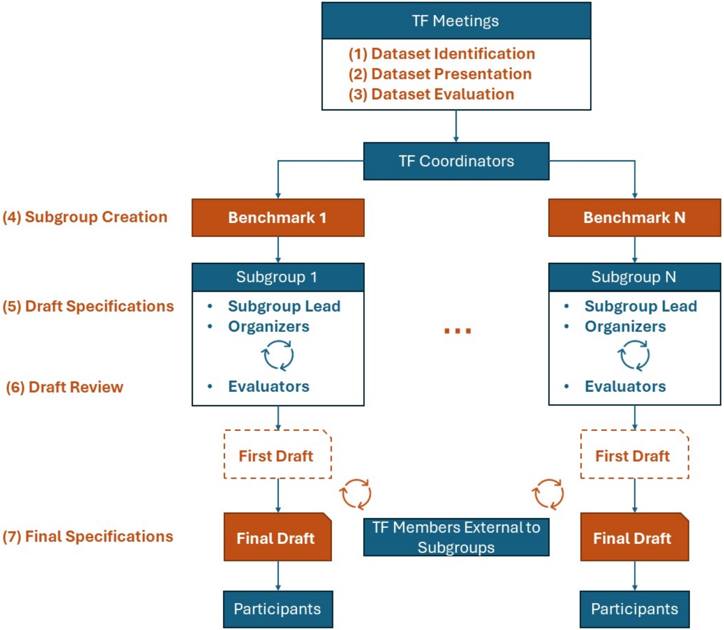
Currently there are 65 members from 28 institutions in 8 countries

- Coordinators
  - TF members who oversee the TF activities together with the NEA secretariat.
- Participants
  - TF members and anyone external to the TF who perform the benchmark exercises and provide their results to the benchmark team.
  - Their main responsibilities are to follow the benchmark guidelines and submit their results within the identified time frame and using the requested format/templates.
- Subgroup Leads/Organizers
  - TF members who lead and manage the development of the specifications for each benchmark exercise.
  - The organizers **cannot** participate in the benchmark since they are the only ones that have access to the blind data.
  - Subgroup leads are also expected to be the organizers of the corresponding benchmark exercise execution.
- Evaluators
  - TF members who are member of the subgroups, involved in designing the tasks, drafting the specifications and providing internal reviews before a wider release of the specifications.
  - The evaluators also **cannot** be participants since they have access to the blind data.
- Reviewers
  - TF members who review the draft specifications and provide comments/suggestions to improve them.
  - The reviewers can also be participants since they do not have access to the blind data.

Benchmark exercise development procedure

1. Dataset identification
   - Identify experimental or numerical datasets to test AI/ML algorithms.
2. Dataset presentation
   - Presentation and discussion of the dataset in the TF status meetings.
3. Dataset evaluation
   - Evaluation of interest and suitability of the dataset for a benchmark exercise.
4. Exercise subgroup creation
   - Identification of 1-2 Leads to oversee a subgroup of members and draft the specification.
5. Exercise draft specifications development
   - The subgroup leads initiate more frequent subgroup meetings to advance the development of the specifications. Updates will be provided in the TF status meetings.
6. Exercise draft specifications review/evaluation
   - Once the exercise specifications are drafted, the Evaluators will need to review and reach an agreement on all details. Then they will be provided to the TF Reviewers.
7. Exercise final specifications
   - Incorporate feedback from Reviewers to finalize the exercise specifications.

TF organization



Outline

List of existing and planned benchmark exercises

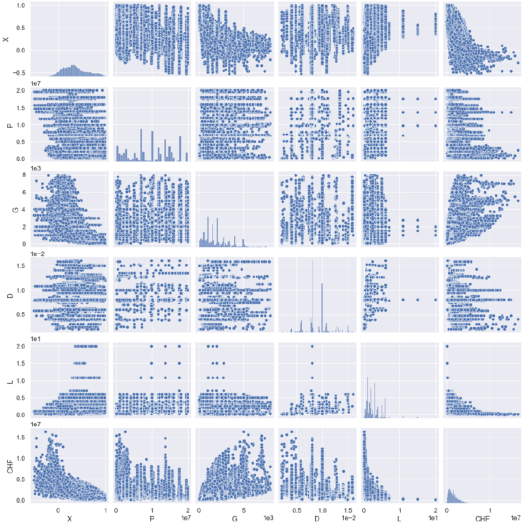
Exercises	Benchmark dataset	AI/ML tasks	Lead	status
1	**Experimental**: NRC Critical Heat Flux (CHF) measurement data	regression, feature selection, UQ, transfer learning	Jean-Marie Le Corre (Westinghouse Electric Sweden)	7/7
2	**Experimental**: Purdue University PUR-1 research reactor operation data	dimensionality reduction, time series regression, anomaly detection	Stylianos Chatzidakis (Purdue University)	6/7
3	**Experimental**: Necsa SAFARI-1 axial neutron flux measurement data	regression, UQ, anomaly detection, clustering	Pavel Bokov and Lesego Moloko (Necsa & NCSU)	5/7
4	**Numerical exercises**: effective Doppler Monte Carlo database, core loading pattern optimization, etc.	regression, UQ, reinforcement learning, etc.	TBD	3/7

Outline

The NRC CHF dataset

Variable	CHF [kW/m^2]	P [kPa]	G [kg/m^2s]	X	D [m]	L [m]
Min Value	50	100	8,2	-0,497	0,002	0,05
Max Value	16339,3	20000	7964	0,999	0,016	20

- The database has 24,579 data points, which is the largest CHF dataset available.
- 62 sources, collected over past 60 years.
- Vertical uniformly heated tubes.
- Large parameter space: coverage well beyond LWR & HWR operational range.
- Base for 2006 CHF lookup table[6].
- Made available by the US NRC[7].

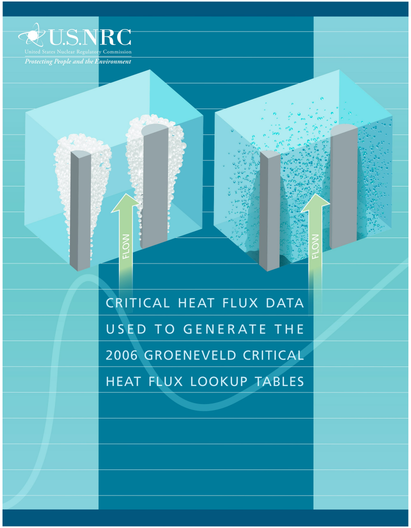


[6] Groeneveld, D. C., Shan, J. Q., Vasić, A. Z., Leung, L. K. H., Durmayaz, A., Yang, J., ... & Tanase, A. (2007). The 2006 CHF look-up table. Nuclear engineering and design, 237(15-17), 1909-1922.

[7] https://www.nrc.gov/reading-rm/doc-collections/nuregs/knowledge/km0011/index.html

The NRC CHF dataset

- The database has 24,579 data points, which is the largest CHF dataset available.
- 62 sources, collected over past 60 years.
- Vertical uniformly heated tubes.
- Large parameter space: coverage well beyond LWR & HWR operational range.
- Base for 2006 CHF lookup table[6].
- Made available by the US NRC[7].



[6] Groeneveld, D. C., Shan, J. Q., Vasić, A. Z., Leung, L. K. H., Durmayaz, A., Yang, J., ... & Tanase, A. (2007). The 2006 CHF look-up table. Nuclear engineering and design, 237(15-17), 1909-1922.

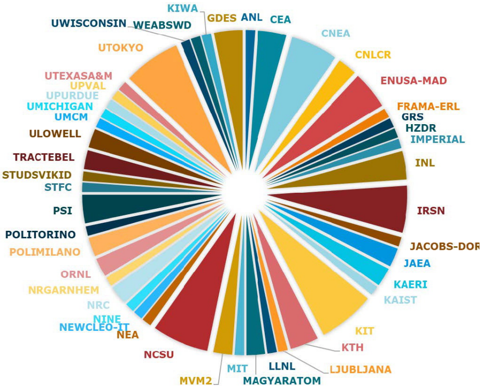
[7] https://www.nrc.gov/reading-rm/doc-collections/nuregs/knowledge/km0011/index.html

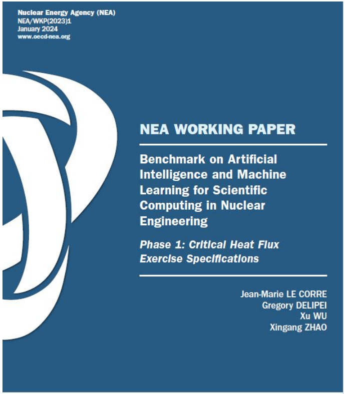
Planned tasks in this benchmark exercise

- The planned AI/ML tasks include:
  - Phase 1 - Regression: to develop ML models that will predict the CHF across various conditions;
  - Phase 1 - Feature analysis: to study what combination of the input features will result the best ML models;
  - Phase 2 - Uncertainty Quantification: to quantify the uncertainty associated with the ML models in order to increase confidence in the predictions;
  - Phase 2 - Transfer Learning: to investigate if the ML models can predict well for unseen geometries (e.g. bundles) and conditions. For this aspect, additional CHF data will be gathered.
- The benchmark organizers team:
  1. Dr. Jean-Marie Le Corre (Westinghouse Electric Sweden, benchmark lead)
  2. Dr. Xingang Zhao (ORNL)
  3. Dr. Xu Wu and Dr. Gregory Delipei (NCSU)
  4. Dr. Oliver Buss (OECD/NEA/WPRS secretariat)

Current status

- The CHF Phase 1 benchmark kick-off was held on October 30th 2023 with more than 90 registered participants from 16 countries and 48 institutions.
- At least 20 benchmark submissions have been confirmed.
- The CHF Phase 1 benchmark specifications were officially published - first NEA Working Paper with number NEA/WKP(2023).





Phase 1 - Timeframe

CHF benchmark introduction at TF meeting	December 2022
Phase 1 draft specification and distribution	May 2023
Presentation at 2023 WPRS Workshops	May 2023
Phase 1 final specifications and distribution	September 2023
Phase 1 online kickoff meeting	October 2023
Phase 1 online Q&A meeting	December 2023
Phase 1 OECD/NEA working paper publication	January 2024
Presentation at 2024 WPRS Workshops	May 2024
Phase 1 submission deadline	August 2024
Phase 1 results draft report and online meeting	December 2024
Presentation at 2025 WPRS Workshops	May 2025

Phase 2 - preliminary timeframe

Phase 2 draft specifications and distribution	May 2024
Presentation at 2024 WPRS Workshops	May 2024
Phase 2 final specifications and distribution	September 2024
Phase 2 online kickoff meeting	October 2024
Phase 2 (fuel bundle) draft specifications and distribution	May 2025
Presentation at 2025 WPRS Workshops	May 2025
Phase 2 submission deadline	August 2025
Phase 2 (fuel bundle) final specifications and distribution	September 2025
Phase 2 (fuel bundle) online kickoff meeting	October 2025
Presentation at 2026 WPRS Workshops	May 2026
Phase 2 (fuel bundle) submission deadline	August 2026

Outline

- The OECD/NEA WPRS Benchmarks on Artificial Intelligence and Machine Learning for Scientific Computing in Nuclear Engineering are first-of-a-kind benchmarks in order to evaluate the performance of AI/ML in multi-physics M&S of nuclear reactor systems.
- These benchmarks will promote international collaboration to further advance and support the development and deployment of AI/ML techniques in diverse nuclear engineering applications.
- If you want to participate in one or more of the benchmark exercises, please contact the TF coordinators or the subgroup leads.
    - TF coordinators:
        - Dr. Xu Wu, xwu27@ncsu.edu
        - Dr. Gregory Delipei, gkdelipe@ncsu.edu
    - CHF benchmark lead
        - Dr. Jean-Marie Le Corre, lecorrjm@westinghouse.com
    - PUR-1 benchmark lead
        - Dr. Stylianos Chatzidakis, schatzid@purdue.edu
    - SARARI-1 benchmark lead
        - Dr. Pavel Bokov, pavel.bokov@necsa.co.za, or pmbokov@ncsu.edu

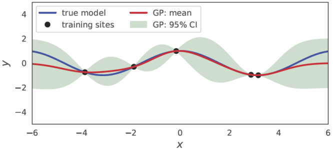
Outline

1. Status of AI/ML in NE
2. OECD/NEA Task Force on AI/ML and Benchmarks
3. Outlook on Near Term Topics

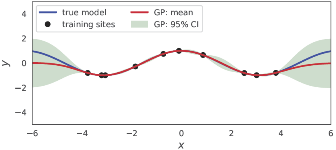
Sources of uncertainties in data-driven Machine Learning models[8]

Data Noise	Data Coverage	Extrapolation	Imperfect Model	Training
noises in training data from either physical simulation models or experiments	few and/or gappy data that has incomplete coverage of training domain	generalization to the extrapolated domains outside of the training domain	ML model architecture is not properly defined, e.g., model is too simple or too complex	random initialization, convergence to local minima, hyper-parameter tuning, etc

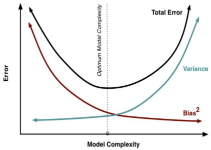
(a)



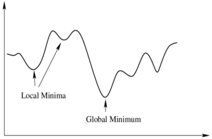
(b)



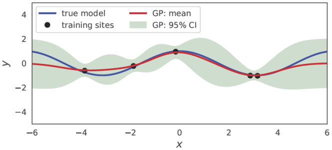
(e)



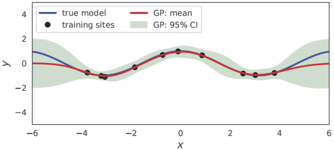
(f)



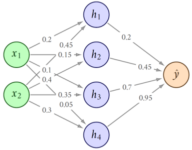
(c)



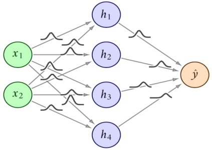
(d)



(g)

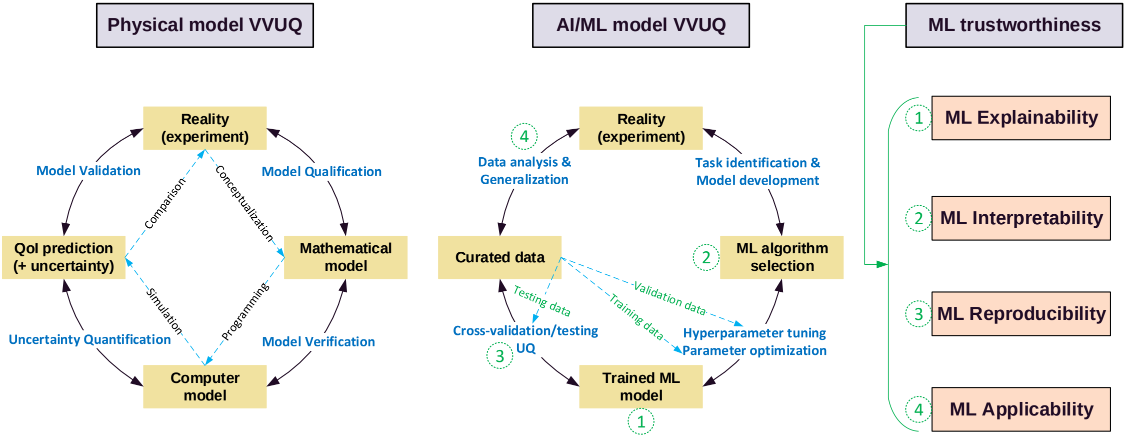


(h)

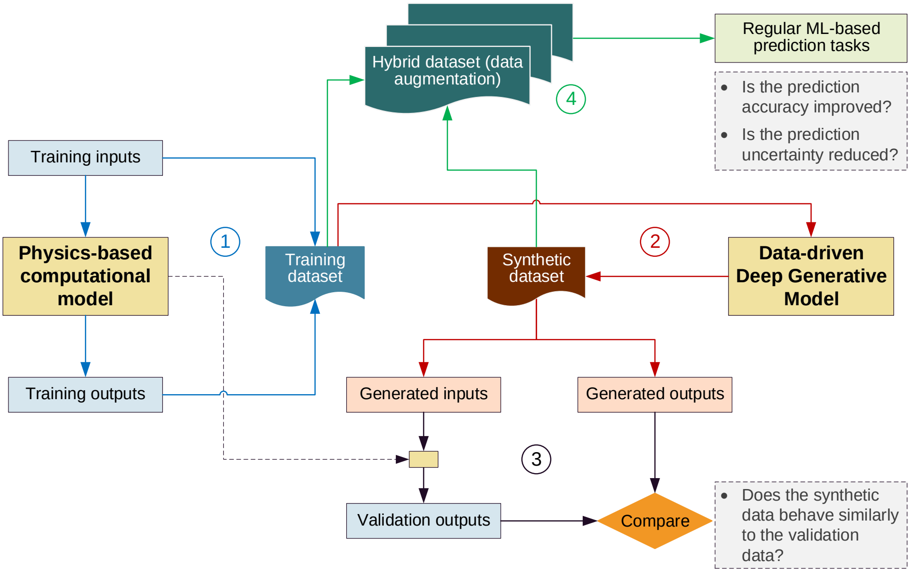


[8] Yaseen, M., & Wu, X. (2023). Quantification of Deep Neural Network Prediction Uncertainties for VVUQ of Machine Learning Models. Nuclear Science and Engineering, 197(5), 947-966.

Rigorous VVUQ is needed for AI/ML in NE applications



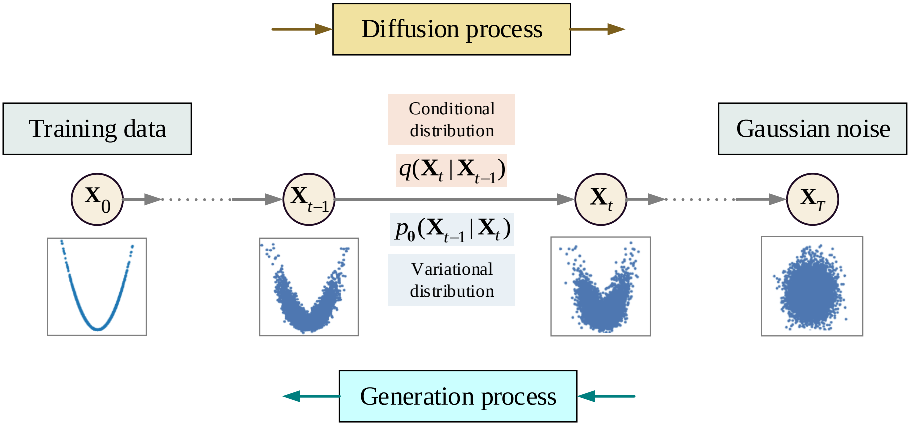
Using Deep Generative Modeling (GANs, VAEs, NFs and DM) for Data Augmentation[9]



[9] Alsafadi, F. and Wu, X. (2023). Deep Generative Modeling-based Data Augmentation with Demonstration using the BFBT Benchmark Void Fraction Datasets. Nuclear Engineering and Design, 415:112712.

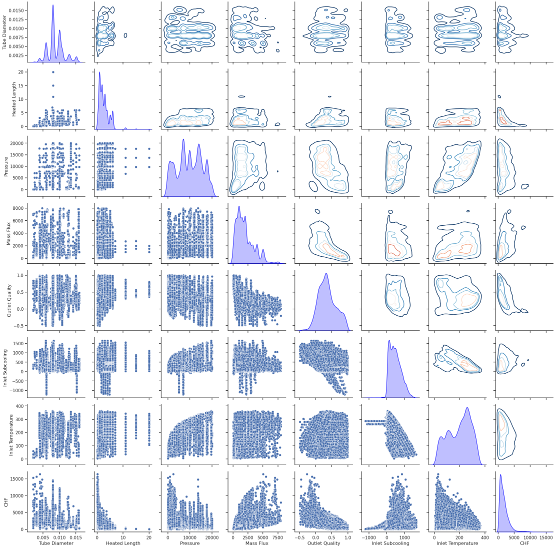
Overview of the Diffusion Model

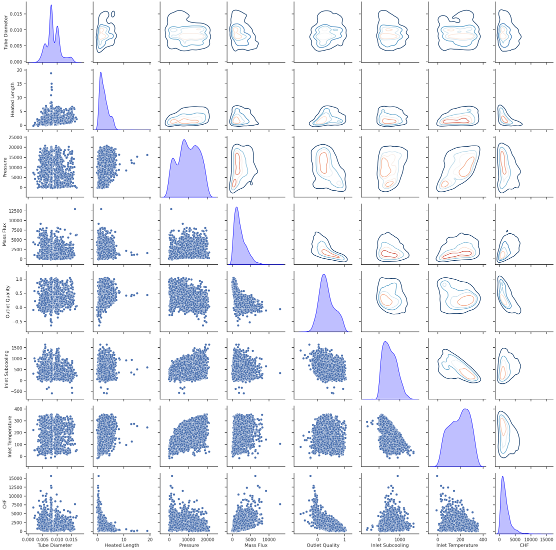
- We have recently investigated the Diffusion Model for scientific data generation.
- OpenAI's Sora[10] is based on Diffusion Model - creating video from text.



[10] https://openai.com/sora

Left: original NRC CHF experimental dataset, right: generated synthetic CHF dataset





NC STATE UNIVERSITY

Thank you for your attention!

Questions and comments?