Reviewer Response for Manuscript NF-104470

Dear Editor,

We thank the referees for their careful consideration and their useful comments. Please find below our detailed responses. We have also made the corresponding changes in the text.

Best Regards,
Petr Mánek (for the authors)

Report of Referee 1

1. The paper presents a comparison of different families of surrogate methods for the calculation of the Tritium Breeding Ration (TBR) in a fusion reactor, claiming major focus to be on Inertial Confinement. In the paper, a "surrogate" method is defined as a cheaper approximation of a much expensive calculation procedure like, for example, a Monte Carlo calculation. Often surrogate methods involve using Artificial Intelligence techniques like, e.g., Artificial Neural Networks.

   The paper compares different families of surrogate methods checking both accuracy and computational complexity, and presents a novel adaptive method for sampling the relevant parameter space, called Quality-Adaptive Surrogate Sampling (QASS), aiming at optimizing the prediction accuracy for a given size of the sampling set or, alternatively, minimizing the number of training points required for a given accuracy. If finally shows that the QASS method is effective at improving the efficiency of a large variety of surrogate methods.

   Although I found the paper interesting to read and acknowledge it can produce a valuable contribution to the development of future fusion technologies, I feel it to be a bit out of the scope of the Nuclear Fusion journal. Numerical techniques are obviously a mission-critical for the development of a scientific and technological task as complex as controlled tehrmonuclear fusion, but a paper with such a strong pitch on computational sciences should be directed more properly, I would suggest, to a journal like, e.g., Journal of Computational Physics or Computer Physics Communications, to which I would recommend to re-direct the submission. I understand, however, that it is up to the editor to take a final decision on the paper acceptability based upon its subject. In the following, I will list my comments as if the paper were to be processed by Nuclear Fusion: I hope they could be useful to improve its quality, whatever journal it will finally be submitted to.

Comments

2. Title: I think the title is a bit misleading. At least to me, it generated first the feeling that the paper were to discuss predictions of the TBR in fusion reactors. I think something more descriptive could be Fast Regression Methods for the Evaluation of the Tritium Breeding Ratio in Fusion Reactors, or something similar.

3. Page 2, line 29, left column: It is stated that the paper presents an empirical surrogate model for the tritium dreeding ratio (TBR) in an inertial confinement fusion (ICF) reactor. This sentence is puzzling to me. If I read correctly the paper later, it compares a family of surrogate methods and develop a strategy for adaptive sampling (QASS). I suspect this is what really the sentence refers to. If it is so, I think the authors should change the sentence, pointing more clearly to QASS.

4. Page 2, line 53, left column: It is stated that the energy distribution for the neutrons considered in the paper is the one called Miur. Actually, a quick serach revealed that this is just a gaussian distribution with certain free parameters. I guess it would be more informative rewording this sentence as a gaussian energy distribution with some user-selected free parameters, re-directing to the relevant literature for more details.

5. Page 2, line 24, right column: While commenting figure 1, it is stated that the geometry of the system were intentionally left adjustable. I suspect that the free parameters are those listed later in Table 1, but this is not stated explicitly. I would appreciate if an explicit comment on which parameters are actually left adjustable could be made at this point.

6. Page 2, line 45, right column: At this point, the paper mentions a novel adaptive sampling procedure suited to this application. I suspect this is what was previously meant (see my comment 2). I think the authors should try to resolve as much as possible this ambiguity between novel surrogate method and novel sampling procedure.

7. Page 3, line 6, left column: It is stated that nuclear fusion technology relies on the production and containement of an extremely hot and dense plasma containing enriched Hydrogen isotopes. While this is certainly true for ICF applications, it is not true for magnetic confinement applications, where the fuel density is approximately 10^{-5} the atmospheric level. It is true that the paper claims to focus on ICF applications, which would make the statement correct. However, in the next few lines it discusses explicitly JET and ITER, which are magnetic fusion experiments. In these few lines there is an oscillation between inertial and magnetic fusion experiments, which make this paper segment ambiguous. I would appreciate if the authors could resolve such ambiguity.

8. Page 3, line 25, left column: It is stated that modern D-T reactors rely on tritium breeding blanket. This sentence seem to be linked to the previous mentioning of JET and ITER. I would appreciate if the authors could be a bit more explicit and precise. It should be noted that JET does not have a blanket, while only test modules will be installed in ITER. It is true that such component is expected to be eventually installed in a reactor. Maybe, an example taken from ICF technology could also be added.

9. Page 3, line 32-39, right column: The authors mention a fast TBR function which takes these same input parameters and approximates the MC model.... It is not clear, to me, what the term function means here. I feel there is again the ambiguity between surrogate methods and sampling technique. Could the authors resolve it?

10. Page 4, line 52, left column: The authors state that they track relative measures that are better suited for comparison between this work and others, because they are invariant... I think it would be better to mention explicitly which ones are the relative metrics and the absolutes (mentioned a few lines earlier).

11. Page 5, line 5, right column: The definition given the for standard deviation does not correspond to the one usually reported in statistics. Could the authors please check it?

12. Page 5, line 51, left column: ... to better characterize the relationship of each family between sample count and the metrics of interest... This sentence is a bit obscure to me. Maybe the authors could re-phrase it?

13. Figure 5: The font-size of the labels is really small and makes it difficult to read. Could the authors increase it?

14. Page 6, lines 25-31, right column: The authors comment the results displayed in figure 3. I understand that the figure was produced by keeping fixed some discrete parameters, but I have no clue which one (or which ones, in case different parameters were kept fixed for different subsets of the data shown). Could the authors explicitly show which one(s)?

15. Page 6, lines 34, right column: It is stated that ERTs, SVMs, and ANNs also achieved satisfactory results with respect to both examined metrics, where I understand the considered metrics are regression performance and prediction time per sample. Actually, the figure suggests that ANNs could be relatively slow. Could the authors elaborate a bit more on that?

16. Page 6, lines 43, right column: It is stated that GBTs, ANNs and ERTs ... are known to be capable of capturing relationships involving mixed feature types that were deliberately withheld in the first experiment. I think it would be appropriate to add a reference to the relevant literature to substantiate this statement.

17. Page 7, lines 30-37, right column: Following both hyperparameter tuning experiments, we conclude that while domain restrictions employed in the first case have proven effective in improving the regression performance of some methods, their performance fluctuates considerably depending on the selected slices. The mentioned fluctuations are not obvious to me. Maybe the authors could elaborate a bit more on this to make them more apparent to the reader.

18. Page 7, lines 54, right column: Commenting the dependence of the performance of the surrogate method families on the dataset size, it is stated that: While such families achieve nearly comparable performance on the largest dataset, in the opposite case tree-based ensemble approaches clearly outperform ANNs. To me, this sentence is

only partially meaningful, because it should be complemented by a comment on the accuracy level that can be attained by whatever methods on the smallest dataset. If it appeared that the achievable accuracy is too low for practical applications, even the speed advantage would be useless. Could the authors add a line of comment/clarification on that?

19. Page 8, right column: The figures inserted in the column are a bit too large. They cover the margin of the left column and some words are chopped. Could the authors fix this?

20. Page 9, lines 8 and 12, left column: The speedups obtained with the surrogate methods are written as $\omega = 6916416\times$ and $\omega = 8659251\times$. I think that writing the acceleration obtained using standard scientific notation 6.92×10^6 and 8.66×10^6 would be far more immediate to read.

21. Page 9, line 41, left column: The authors introduce a sinusoidal theory as an alternative to the expensive MC Paramak model used to assess the efficiency of the surrogate methods and sampling techniques they analyze. They qualify this theory as robust. It is by far not clear to be what the word robust means in this context. Could the authors clarify this?

22. Page 10, line 4: The best-performing method(s) under each are highlighted in bolt I do not fully understand this sentence. Perhaps a word is missing (under each ... what ... ?)

23. Page 10, figure 7 and 8: They are too large and covers some letters of the opposite left column. I think they should be resized.

24. Page 11, line 7, left column: An increase in initial sample size was observed to also resolve precision in these smooth regions... I do not understand what the wording resolve precision means in this context. Maybe the authors could elaborate a bit more on that?

25. Page 11, line 26-31, left column: The authors claim they developed fast and high-quality surrogates for the Paramak TBR model. I suspect that here again there is an ambiguity between surrogate models and sampling strategies. If the authors really developed a novel surrogate model, this is not clear to me. Could the authors elaborate more on that?

26. Page 12, line 32, left column: When following the url reported in the citation, I reach a publication with title The components of the Wired Spanning Forest are recurrent, which looks different from the declared one. Maybe is the link wrong?

Report of Referee 2

1. Highly interesting work on empirical surrogate modelling for an integral 3D Monte-Carlo radiation transport problem (TBR).

   Comments

2. Introduction: use capital letters for FENDL and JEFF projects; ENDF/B-VII.1.

3. Introduction: I am not fully convinced that VR is a competitor of your work or perhaps in other applications. The concept is very different, the objective is similar: to achieve significant speedups.

4. 1.1: The acronym ITER should be used without the abandoned explanation. Just say "ITER".

5. 1.1: 2nd paragraph starts with "instead". However, the tritium self-sufficiency is a paramount requirement for viable nuclear fusion energy production (DT reaction). Therefore the tritium breeding is an inherent feature of DT fusion reactors; there is no way to rely, even partially, on external supply, except of start-up inventories.

6. 1.1: TBR definition is ok, if "consumption" is well understood. Better to use the standard definiton tritium produced per source neutron (or tritium fused in plasma).

7. 1.1: Table 1: Thickness in mm; Armour material I suppose is W.

8. 1.1: It seems to me that the accessible domains could be narrowed down for a reasonable design space. Would this impact the assessment made in the paper? This obviously applies most prominently to parameters like armour fraction and blanket thickness.

9. 2: Table 2: it might be useful to briefly introduce the rationale of hyperparameter sets (how they are built from the Paramak input). What could be said about the size, e.g. using only 1-2 parameters vs. 11?

10. 2.2: Regarding QASS, I am confused about the objective to maximize surrogate eorr (page 6). Is this a typo?

11. 2.2: The CDM is not introduced through ref. 26. I would anyway suggest to describe the core of the metric in your article.

12. 3.2: what does "most robust" mean for the selected TBR toy theory? Is "robustness" a requirement for the functionality or success of QASS?

13. 3.2: Fig. 7/8: explain E-MAE in the article and how it relates to the introduced MAE.

Report of Board of Editors

1. The reviewer believes to publish this paper in the Nuclear Fusion, the authors should interpret or correlate finding/result/significance with the fusion terminology/knowledge/practice. In this case, what is the final design composition (the set of input parameters listed in Table 1) derived from the proposed methodology that would provide an adequate TBR, and how are the derived design compositions compared with the available designs? Without such a comparison, it is hard to evaluate the contribution of the paper toward the advancement of fusion nuclear science and technology (FNST). Furthermore, even TBR is a critical parameter in the design, the blanket design has to satisfy other constraints such as temperature, stress, MHD pressure drop, etc. How does the proposed method incorporate additional yet necessary constraints in searching for a required TBR?

   The goal of the proposed methodology is not clearly stated. If the goal is only to produce a fast TBR function as stated below "In our work, we set out to produce a fast TBR function which takes these same input parameters and approximates the MC model used in Paramak with the greatest achievable regression performance, while also minimizing the required quantity of expensive samples," the reviewer would consider this work is in its primitive stage of the application of artificial intelligence/machine learning to FNST, and a too narrow fusion relevant scope to be published in Nuclear Fusion. More thoughts should be given to the identifications of the space/critical design needs from an integrated perspective (versus a single parameter such as TBR), where AI/the proposed methodology can help ensure tritium fuel self-sufficiency for the fusion energy development.