9 June 2025 ·

Qeios

Peer Review

Review of: "Development of an Uncertainty-Aware Equation of State for Gold"

Peter Hatfield[1]

1. Department of Physics, University of Oxford, United Kingdom

Very interesting work. I particularly enjoy it for using Gaussian Processes rather than NNs (I think each has its place) and for, in my opinion, importantly honing in on trying to separate the multiple different sources of uncertainty. No major comments, apart from echoing many of the comments from other reviewers with which I largely agreed, but a few small things I would add:

There is a small body of papers that I thought might be worth looking at or for a comparison, mainly the sparse Gaussian Process code GPz (originally developed for the problem of photometric redshifts in astronomy) and a few applications of the algorithm to HEDP:

- Almosallam developed in his thesis some theory and an excellent code considering several of the issues in this paper about how to have GPs quantify multiple different sources of uncertainty. His framework explicitly quantifies uncertainty from i) uncertainty due to error bars on input data, ii) uncertainty due to sparsity of data in a part of parameter space, and iii) intrinsic scatter in output (with iii) possibly having the subtle variation of uncertainty in measurement in output versus true variation). Your Error in Variables is conceptually similar, in my understanding, to the \gamma in the Almosallam framework (he also has \beta for noise on the output parameter, and \nu for uncertainty from lack of data), so it might be interesting to compare how similar his work is. The main papers are arXiv:1505.05489 and arXiv:1604.03593 (potentially also his thesis and other papers written with Roberts). You also cite scaling concerns for higher dimensions with the traditional O(N^3); GPz uses a sparse framework and scales with O(N) and has been used extensively in greater than 10D with training data in the 10-100s of thousands with relatively standard computing facilities, so it might be of interest for that reason too.

 doi.org/10.32388/OQESZJ

- Hatfield et al. (apologies for the self-cite) wrote a few papers applying GPz to HEDP problems to try and use the advantages of GPs. Probably most relevant to this work is arXiv:1910.08410, in which I used GPz in a similar way to your paper but for ICF yields rather than EoS. In that paper, I tried to separate the uncertainty out into uncertainty from lack of data in a part of parameter space versus uncertainty due to uncertainty on output (right-hand panel of figure 4). The paper also more broadly shows GPs scaling well and, in some contexts, outperforming NNs. I add in the third source of uncertainty, Error from Input variables, in these two astronomy papers (arXiv:2009.01952 and arXiv:2206.00748) for three components to the GP error in total, and also discuss different ways of combining experiment and theory, e.g., Figure 2 in the second of those papers, where I map out different parts of parameter space and identify in which parts theory performs well and which parts data-driven ML models are better.

Other minor points:

- Would it be possible to get a sentence or two describing what your vision might be for the next stages in this line of inquiry? Are you imagining a rad-hydro code like Hyades might sample from the resulting model when doing predictions to give better uncertainties on predictions of ICF simulations?

- I would find it helpful to understand a little better how DFT theoretical predictions and experimental data could be optimally combined in the future, e.g., smoothly transitioning between a part of parameter space where you are regressing on data to where you are regressing on theory?

- Not a major point, but in the Abstract and Conclusion, you largely pitch this as a data science paper, with the statistical elements of the work as being the novel parts. Section III essentially just focuses on the physics of the DFT, etc. - I think you do need it in there somewhere, but many readers might be happy to trust that the physics part was done well, with only a small number of people wanting to look up the specifics, so maybe some of this could go in an Appendix.

- I think it might be helpful to explain how you hoped to draw conclusions from the EIV. For example, would you be able to show from work like this that errors in our ability to measure pressure accurately were the main thing holding back EoS, and thus motivate improved instrumentation focusing on that, or alternatively that it is the number of data points rather than measurement accuracy, in which case you would want more shots (or whatever the conclusions might be). I guess basically draw out how splitting the uncertainty up leads to greater insight into how best to reduce the overall uncertainty.

Declarations

doi.org/10.32388/OQESZJ

Potential competing interests: No potential competing interests to declare.

doi.org/10.32388/OQESZJ