f3dasm: Framework for Data-Driven Design and Analysis of Structures and Materials

M. P. van der Schelling [1], **B. P. Ferreira** [2], and **M. A. Bessa** [2]¶

1 Materials Science & Engineering, Delft University of Technology, the Netherlands **2** School of Engineering, Brown University, United States of America ¶ Corresponding author

Summary

`f3dasm` (Framework for Data-driven Design and Analysis of Structures and Materials) is a Python project that provides a general and user-friendly data-driven framework for researchers and practitioners working on the design and analysis of materials and structures. The package aims to streamline the data-driven process and make it easier to replicate research articles in this field, as well as share new work with the community.

Figure 1: Logo of `f3dasm`.

Statement of need

In the last decades, advancements in computational resources have accelerated novel inverse design approaches for structures and materials. In particular, data-driven methods leveraging machine learning techniques play a major role in shaping our design processes today.

Constructing a large material response database poses practical challenges, such as proper data management, efficient parallel computing, and integration with third-party software. Because most applied fields remain conservative when it comes to openly sharing databases and software, a lot of research time is instead being allocated to implement common procedures that would be otherwise readily available. This lack of shared practices also leads to compatibility issues for benchmarking and replication of results by violating the FAIR principles.

In this work we introduce an interface for researchers and practitioners working on the design and analysis of materials and structures. The package is called `f3dasm` (Framework for Data-driven Design and Analysis of Structures and Materials). This work generalizes the original closed-source framework proposed by Bessa and co-workers (Bessa et al., 2017), making it more flexible and adaptable to different applications, namely by allowing the integration of different choices of software packages needed in the different steps of the data-driven process:

- **design of experiments**, in which input variables describing the microstructure, properties and external conditions of the system are determined and sampled;

- **data generation**, typically through computational analyses, resulting in the creation of a material response database (Ferreira et al., 2023);
- **machine learning**, in which a surrogate model is trained to fit experimental findings;
- **optimization**, where we try to iteratively improve the design.

Figure 2 provides an illustration of the stages in the data-driven process.
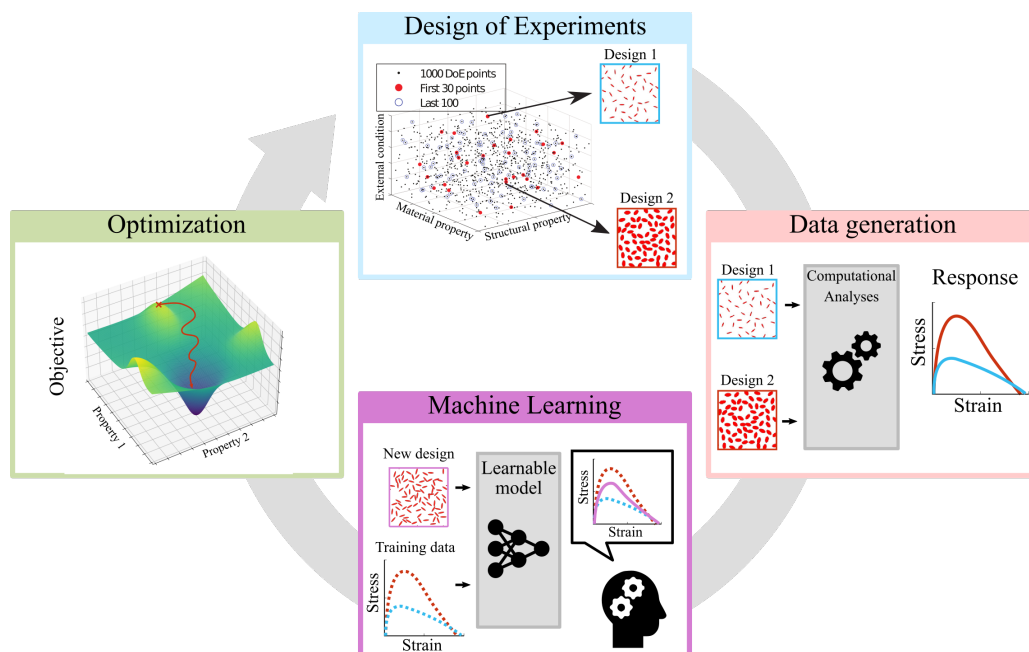

Figure 2: Illustration of the f3dasm data-driven process.

f3dasm is an open-source Python package compatible with Python 3.8 or later. The library includes a suite of benchmark functions, optimization algorithms, and sampling strategies to serve as default implementations. Furthermore, f3dasm offers automatic data management for experiments, easy integration with high-performance computing systems, and compatibility with the hydra configuration manager. Comprehensive online documentation is also available to assist users and developers of the framework.

In a similar scope, it is worth mentioning the projects simmate (Sundberg et al., 2022) and strucscan, as they provide tools for the management of materials science simulation and databases. However, these projects focus on the generation and retrieval of materials properties and do not include machine learning or optimization interfaces. In recent years, numerous optimization frameworks have been developed to facilitate data-driven design. Optuna is a hyperparameter optimization framework that combines a variety of optimization algorithms with dynamically constructed search space (Akiba et al., 2019) and pygmo provides unified interfaces for parallel global optimization (Biscani & Izzo, 2020). Interfaces to these and many other optimization frameworks have been integrated into a separate package f3dasm_optimize, and can be used in conjunction with f3dasm.

Acknowledgements

References

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2623–2631. https://doi.org/10.1145/3292500.3330701

Bessa, M., Bostanabad, R., Liu, Z., Hu, A., Apley, D. W., Brinson, C., Chen, W., & Liu, W. K. (2017). A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality. *Computer Methods in Applied Mechanics and Engineering*, *320*(April), 633–667. https://doi.org/10.1016/j.cma.2017.03.037

Biscani, F., & Izzo, D. (2020). A parallel global multiobjective framework for optimization: pagmo. *Journal of Open Source Software*, *5*(53), 2338. https://doi.org/10.21105/joss.02338

Ferreira, B. P., Pires, F. M. A., & Bessa, M. A. (2023). CRATE: A python package to perform fast material simulations. *Journal of Open Source Software*, *8*(87), 5594. https://doi.org/10.21105/joss.05594

Sundberg, J. D., Benjamin, S. S., McRae, L. M., & Warren, S. C. (2022). Simmate: A framework for materials science. *Journal of Open Source Software*, *7*(75), 4364. https://doi.org/10.21105/joss.04364