DION HÄFNER

PROGRAMMING SCIENTIST | SCIENTIFIC PROGRAMMER

Malmö, Sweden mail@dionhaefner.de dionhaefner.github.io dionhaefner

I am a Physicist specialized on all things scientific computing, with industry experience in data processing and machine learning.

My main research interest is "human learning": How can we build machines that teach us something about our physical world?

Education

Dec 2018 – present **PhD in Physical Oceanography @ University of Copenhagen**
In my PhD project at the Niels Bohr Institute (Prof. Markus Jochum) I inferred the relative importance of hypothesized causes of extreme ocean waves ("rogue waves") in the real world. To this end, I analyzed over 1 Terabyte of observational data with data mining and probabilistic machine learning to extract robust, causal insights.

Sep 2021 – Dec 2021 **Research visit @ University of Valencia**
During my 3-month research visit to the Image Processing Lab (Prof. Gustau Camps-Valls) I investigated how machine learning can help humans to understand climate models.

Sep 2015 – Dec 2016 **External Master's thesis in Physical Oceanography @ University of Copenhagen**
Working with state-of-the-art climate models, I learned how to process and analyze large quantities of data, and became familiar with geoscientific modeling and high-performance computing.

Sep 2011 – Dec 2016 **BSc and MSc in Physics @ Heidelberg University**
GPA of 1.3 and 1.2, respectively ("very good"). Exchange year at KTH Stockholm in 2014. Specialization on computational physics.

Professional Experience

Apr 2022 – present **Research Engineer @ Pasteur Labs**
I build efficient, scalable software for scientific advances in all areas of simulation intelligence (SI). This includes differentiable programming, earth systems simulation, physics-infused machine learning, probabilistic machine learning and computation, simulation based inference, and causal machine learning.

Sep 2017 – Dec 2018 **Software development specialist @ DHI GRAS**
At DHI GRAS, I built robust data pipelines, powerful statistical tools, and optimized remote sensing workflows.

Jan 2017 – Aug 2017 **Scientific assistant @ Niels Bohr Institute**
I developed the first prototype of Veros, a high-performance ocean model in Python.

Jun 2014 – Dec 2016 **Research assistant @ Institute of Environmental Physics**
Working in a small team of developers on a numerical software suite for soil water flow (discontinuous Galerkin PDE solver), in C++ and Python.

Software Projects

Veros — A high-performance ocean model in pure Python

I am the main developer and maintainer of Veros, a full-fledged primitive equation ocean model capable of accurate, realistic simulations of the global ocean. It leverages the JAX library for state-of-the-art performance on CPU and GPU clusters.

https://github.com/team-ocean/veros

Terracotta — A light-weight geospatial raster tile server

Terracotta is a cloud-ready raster tile server, leveraging the cloud-optimized GeoTiff format and a modern geospatial Python stack.

https://github.com/DHI-GRAS/terracotta

mpi4jax — Zero-copy MPI communication of JAX arrays

mpi4jax enables zero-copy, multi-host communication of JAX arrays, even from compiled code and from GPU memory.

https://github.com/mpi4jax/mpi4jax

(see all on **GitHub**)

Programming and Technology

ML frameworks I am familiar with modern machine learning workflows and have good knowledge of scikit-learn, Tensorflow / Keras, PyMC, and JAX.

Python I have both deep and broad experience within the Python ecosystem, especially concerning (but not limited to) **data analysis, machine learning, visualization, and scientific computing**. I love working with the modern scientific Python stack and am well-versed with NumPy, SciPy, matplotlib, xarray, JAX, Numba, and Pandas.

Other languages Basic knowledge of object-oriented programming in C++, including debugging applications with GDB. Elementary knowledge of OpenCL / CUDA.

Tools Experience with tools handling version control (git), documentation (Sphinx, Doxygen), build systems (CMake), deployment (Docker), testing (pytest), GUI (Qt), continuous integration (Travis CI / Github Actions), and typesetting (LaTeX).

I am comfortable working in all major operating systems, and am familiar with basic server administration tasks, including cloud providers like AWS and GCP.

Other Skills & Interests

- Strong **mathematical and analytical skills**, and an affection for data.
- Good theoretical foundation of **applied mathematics and scientific computing** (including ODE / PDE solvers, numerical optimization, signal processing, and automatic differentiation).

» A knack for probabilistic reasoning and **Bayesian data analysis**: I like to make my assumptions and uncertainties explicit.

» I am passionate about **open-source software development**, and have contributed code to several big projects on GitHub.

» A special interest in **effective communication** through writing, oral presentations, and data visualization. I take the quality of my publications seriously, and love to present my work.

» **Languages**: German (native), English (fully proficient), Swedish (proficient), Danish (elementary).

Teaching

Courses 3 Teaching Assistant (TA) assignments in Physics MSc courses at the Niels Bohr Institute. Nominated for biannual TA prize in 2021.

Theses Co-supervisor to 1 BSc student and 1 MSc student at NBI. Closely involved in 5 MSc projects from other departments / universities.

Invited Talks

» (2022) *"JAX of all trades — GPU-accelerated ocean modelling in Python"*. Keynote presentation @ DRAKKAR ocean modelling workshop (virtual).

Awards & Honors

» (2021) Virtual Outstanding Student and PhD candidate Presentation (vOSPP) by European Geophysical Union.

Peer-reviewed Publications

[1] **Häfner, D.**, Nuterman, R., Jochum, M., "Fast, Cheap, and Turbulent—Global Ocean Modeling With GPU Acceleration in Python". In: *Journal of Advances in Modeling Earth Systems* 13.12 (Dec. 2021). DOI: 10.1029/2021ms002717. URL: https://doi.org/10.1029%2F2021ms002717.

[2] **Häfner, D.**, Vicentini, F., "mpi4jax: Zero-copy MPI communication of JAX arrays". In: *Journal of Open Source Software* 6.65 (Sept. 2021), p. 3419. DOI: 10.21105/joss.03419. URL: https://doi.org/10.21105%2Fjoss.03419.

[3] **Häfner, D.**, Gemmrich, J., Jochum, M., "FOWD: A Free Ocean Wave Dataset for Data Mining and Machine Learning". In: *Journal of Atmospheric and Oceanic Technology* (May 2021). DOI: 10.1175/jtech-d-20-0185.1. URL: https://doi.org/10.1175%2Fjtech-d-20-0185.1.

[4] **Häfner, D.**, Gemmrich, J., Jochum, M., "Real-world rogue wave probabilities". In: *Scientific Reports* 11.1 (May 2021). DOI: 10.1038/s41598-021-89359-1. URL: https://doi.org/10.1038%2Fs41598-021-89359-1.

[5] Riedel, L., Ríos, S. D., **Häfner, D.**, Klein, O., "DORiE: A Discontinuous Galerkin Solver for Soil Water Flow and Passive Solute Transport Based on DUNE". In: *Journal of Open Source*

Software 5.52 (Aug. 2020), p. 2313. DOI: 10.21105/joss.02313. URL: https://doi.org/10.21105%2Fjoss.02313.

[6] **Häfner, D.**, Jacobsen, R. L., Eden, C., Kristensen, M. R. B., Jochum, M., Nuterman, R., Vinter, B., "Veros v0.1 – a fast and versatile ocean simulator in pure Python". In: *Geoscientific Model Development* 11.8 (Aug. 2018), pp. 3299–3312. DOI: 10.5194/gmd-11-3299-2018. URL: https://doi.org/10.5194%2Fgmd-11-3299-2018.

(see also **Google Scholar**)

References

Markus Jochum
Professor, Niels Bohr Institute, University of Copenhagen, Denmark
mjochum@nbi.ku.dk

Gustau Camps-Valls
Professor, Image Processing Lab, University of Valencia, Spain
gustau.camps@uv.es

Johannes Gemmrich
Research Scientist, Physics & Astronomy, University of Victoria, Canada
gemmrich@uvic.ca