natESM Newsletter

ISSUE 8 / JANUARY / 2025

Contents

deep dive Open Development

How can we ensure that natESM remains an open and community-driven system? In this issue's deep dive on the next two pages, we explore possible principles, challenges, and opportunities of Open Development and its significance for the natESM community.

Check out our education site

We're happy to announce the launch of the →natESM education website, developed by the Training Working Group under the leadership of Birgit Hassler (DLR-PA). This platform is a treasure trove of resources for Earth system modelers at all levels, from beginners to experts.

The site includes:

- Tutorials for getting started with Earth system modeling
- Hands-on exercises to deepen technical skills
- Advanced materials for exploring cutting-edge methods

DWD ICON course in summer

This summer, natESM proudly supports the DWD Academic ICON Course—a training program for newcomers to numerical modeling. Participants will gain essential skills to effectively use the ICON model. Applications open this spring, but you can learn more about the course and its objectives →here. Stay tuned for updates!

Join us for the annual community workshop!

📅 When: February 18–19, 2025
📍 Where: Hotel Aquino, Berlin
📝 **Register by: January 24, 2025**

This year's workshop is again your chance to help design the future of natESM. While the project was initially set to conclude this February, a cost-neutral extension until August 31, 2025, gives us extra time to reflect, plan, and collaborate on a potential second phase of natESM.

The agenda features discussions on sprint outcomes, keynotes on AI and ICON atmosphere, and breakout sessions to shape natESM's future. You'll also have the opportunity to meet the scientists and RSEs behind the latest sprints, who will present their results and share insights into their experiences with collaboration, challenges, and achievements.

👉 Register here before January 24. You can find the agenda here.

Refining the sprint process

To better support the natESM community, we have refined the sprint process in our agile strategy with a clear focus on achieving long-term goals. The updated sprint process emphasizes:

- Stronger collaboration between scientists and RSEs during and after sprints
- Alignment with institutional workflows to ensure sustainability of results
- Building a foundation for a potential second phase of natESM

These changes aim to make our processes more effective and to ensure that every sprint leaves a lasting impact.

👉 The updated strategy is available →here, with the sprint process starting at the bottom of page 8.

deep dive

personal stories from natESM trailblazers

Building bridges through Open Development

Interview by Iris Ehlert

Roland Potthast is Director of the Division for Numerical Weather Prediction at DWD and Professor for Applied Mathematics at the University of Reading.

Openness is essential to natESM's vision of building a collaborative platform, but in the rush of our daily work, we rarely take the time to reflect on what "openness" really means or what it entails for collaboration. For this deep dive, I spoke with two members of our community who, through their experiences during the process of making the ICON code open source, have gained valuable insights into the challenges and opportunities of Open Source and Open Development. Their reflections shed light on what it takes to build bridges between institutions and people while fostering a culture of openness. We talked for over an hour, and here I can only share a small part of our exchange. If you want to dive deeper, you'll have the chance to exchange ideas with them in person during our community workshop in Berlin next month.

Hendryk Bockelmann is head of the Application Support Department at DKRZ and PI of the natESM project.

natESM: When you think about natESM, what do "Open Source" and "Open Development" personally mean to you?

Hendryk: Open Source means that the code is freely accessible. It ensures that users can always access at least the latest version of the code and that in the best case all developments are publicly visible. In the natESM context, this is crucial because our goal is to enable every scientist to benefit from the advancements that our Research Software Engineers (RSEs) implement.

Open Development is the next step: it involves making the development process itself open. This means anyone can offer contributions, provided their contributions are meaningful and adhere to the established standards. This approach fosters collaboration and exchange—a central goal of natESM.

Roland: While Open Source focuses on free access to the code, Open Development describes the interaction and collaboration within the community. It ensures not only that the code is accessible but also that the entire development process is transparent and open. For natESM, this is particularly important as we aim to create not just a product, but a platform supported by the community. It's about enabling both users and developers to contribute to success.

In your opinion, why is embracing Open Source such a vital step for natESM's growth?

Roland: Science is built on openness. In research, we expect results to be published and made accessible—that's essentially the principle of "Open Science." Why should it be different with software? Software is often a central part of research, and if it isn't freely accessible, results can't be reproduced. Open Source is therefore essential to ensuring the integrity of scientific work.

Licenses are also an important aspect. Permissive licenses, like the BSD-3-Clause, allow our software to be used both in academia and by industry. This is critical for collaborations with companies, such as those developing compilers.

Hendryk: I'd like to add that Open Source alone isn't sufficient if the code is available but cannot be modified. That's more akin to "Open Access" than "Open Source". Science thrives on the evolution of ideas. It's vital that not only the results but also the tools are open and adaptable.

What do you see as the biggest hurdles or growing pains in making Open Development work for natESM?

Hendryk: Open Development requires clear structures. We need tests, documentation, and infrastructure that enable efficient contribution. This is time- and resource-intensive. A significant issue is that most developers in science are mainly researchers, leaving little time to focus entirely on the development process.

Additionally, there's often a lack of technical support and governance structures that allow external contributions. Currently, this mostly works through personal contacts and partnerships, but that's not sustainable.

Roland: I agree. A major goal for natESM is achieving modularity. This means designing our systems so that different teams or institutions can contribute meaningfully without requiring central coordination. This requires well-documented modules and clear interfaces. However, that's easier said than done. It demands experienced developers capable of building such systems and providing long-term support.

» In my experience, Open Development has the power to break down silos. It's not just about code; it's about creating a framework where teams align their efforts and discover unexpected synergies. «

Hendryk: Moreover, Open Development has a cultural component. It requires a mindset of openness and mutual respect. We need to actively work on creating a community where contributions are not only technically reviewed but also valued.

Roland: Another point to consider is the balance between openness and quality. While openness brings many advantages, it also requires additional resources for code reviews, feedback, and onboarding new members. Balancing this with the available capacities can be challenging. If you have a life system, which is used daily and monthly for operational weather and climate forecasts and projections, keeping its quality fully intact is of highest importance.

From your experience, how does Open Development change the way teams collaborate within natESM?

Roland: Collaboration is the backbone of Open Development. One aspect we've noticed is that it helps break down silos within the scientific community. For instance, when two teams working on seemingly unrelated components discover overlaps or synergies, Open Development provides a framework for them to collaborate effectively.

Hendryk: Exactly. And it's not just about technical collaboration but also about sharing knowledge. When teams document their processes and make them available, it creates a resource that others can learn from. This kind of shared learning accelerates progress across the board.

Roland: A recent example in natESM was when one team working on climate simulations collaborated with another focusing on land-use modeling. Open Development enabled them to align their efforts, resulting in a module that serves both areas. Without this open approach, such synergies would have been harder to realize.

Hendryk: That's a great point. It shows how Open Development isn't just a technical challenge but also a cultural shift. It encourages people to think beyond their immediate projects and consider the bigger picture.

Looking ahead, how do you imagine Open Development shaping natESM's future?

Hendryk: I think the first step is modularizing our models and codes. By clearly defining and documenting modules (sometimes also called building blocks), it becomes easier for different teams to work in parallel. Additionally, we need more resources for testing and documentation.

Currently, the focus is often on developing new features, but without a stable foundation, this won't be sustainable in the long run. Another crucial point is involving the community. We need to ensure that all stakeholders are heard, and their needs considered. This requires ongoing dialogue and open communication.

Roland: I agree. Modularity is key, but we also need to remain flexible. Overly rigid structures can stifle creativity. Iterative development, where we continuously gather feedback and adjust, is, in my opinion, the best approach. In the long term, we should also consider how to involve more developers in the community. This could include special training or mentoring programs. We should also think about how to integrate new members into the natESM community without overwhelming them.

Hendryk: Communication plays a vital role here. It's important to clearly communicate the value of Open Development. Often, Open Source is perceived as free but demanding. We need to show that it's worth it—in terms of scientific progress and personal development.

» I see Open Development as more than the technical side—it's about culture and communication. By working together as a community, we can create something that truly makes a difference. «

Roland: Additionally, we should consider strengthening long-term collaborations with other projects. Interdisciplinary cooperation can help bring in new perspectives and address challenges more effectively.

As we wrap up, is there something you'd like to share with the community?

Hendryk: Open Source and Open Development are not just technical challenges. They are about culture and communication. We need to ensure that everyone in the community feels welcome and that their contributions are appreciated.

Roland: For me, it's clear: Open Development is the way forward, not just for natESM but for science as a whole. We have a great opportunity to build something sustainable and innovative here. I'm optimistic that we can achieve this together.

Hendryk: Yes, it's not always easy, but the effort is worth it. By working together as a community, we can create something that goes beyond natESM and truly makes a difference.

Roland: Exactly. The challenge lies in fostering a culture centered on openness, respect, and collaboration. I'm confident that we can achieve this if we all pull together.

Sprint status

Sprint title	Inst.	Service description
ICON-ART Finished	KIT	Analysis of ART code for GPU porting → Sprint report
ICON-mHM-YAC Finished	UFZ	Online coupling mHM into ICON using YAC → Sprint report
FESOM Finished	AWI	Port FESOM 2.1 to JUWELS booster and Levante-GPU → Sprint report
ParFlow Finished	FZJ	Port ParFlow to AMD GPUs, Performance Analysis → Sprint report
MESSy Finished	FZJ	Optimize the data transfers between host (CPU) and device (GPU) → Sprint report
ESMValTool Finished	DLR-PA	Updating remaining non-lazy preprocessor functions to be memory efficient → Sprint report
HAMOCC Finished	MPI-M	Concurrent HAMOCC on GPU → Sprint report
MESSy-ComIn Finished	DLR-PA	Couple MESSy to ICON via the ICON Community Interface → Sprint report
LAGOOn Finished	FZJ	Develop concept and provide first implementation of Lagrangian-transport-modeling framework → Sprint report just finished!
IQ Finished	MPI-BGC	Stepwise port of IQ code to GPUs based on established workflow followed for ICON-GPU implementation with OpenACC → Sprint report just finished!
modLSMcoup Reporting	FZJ	Develop proof-of-concept for modular coupling of land surface and implement YAC coupler in ICON-eCLM coupling
CLEO Finished	MPI-M	Coupling CLEO to ICON with YAC → Sprint report just finished!
PALM Running	Uni Hannover	Porting those of the PALM modules related to urban processes (especially radiation) to GPUs
MESSy-ComIn2 Running	DLR-PA	ComIn integration time loop
PDAF2GPU Running	AWI	Porting PDAF to GPUs
MESSy-IMPORT Running	FZJ	Revise the data import function of the Modular Earth Sub-model System (MESSy) for ICON/MESSy
MESSy-GPU port Running	FZJ	Enabling ICON/MESSy for GPUs with additional GPU porting of MESSy submodels This sprint is being conducted by our associated team member in the esiwace project.
PISM-AsyncIO Waiting	MPI-GEA	Resolve the issue with the I/O library for asynchronous output

Additional information from the sprints, beyond what is covered in the sprint reports, is available in our GitLab wiki.

REMINDER: Do you have a model, module, or tool you'd like to share with the natESM community? You can use our **natESM GitLab system** to showcase your work and increase its visibility. If you have something to share, fill out →this form, and we'll take care of the rest!