Newsletter

SPECIAL EDITION: COMMUNITY WORKSHOP / FEBRUARY 2025

Slides from all presentations are available on **our website**.

Contents

Welcome to our special edition on the natESM community workshop! Our Co-Chair, Anja Schmidt, kicked off the event with an update on the progress and our future plans for Earth system modeling in Germany.

Progress and process improvements

- **Sprint checks and applications:** We have implemented a new sprint-check process—a preliminary evaluation that determines if a piece of code is technically ready for integration into natESM. This step not only streamlines the review process by filtering out proposals that need further refinement but also guides developers in preparing their applications for full sprints.
- **Weekly meetings:** Regular weekly meetings have been established to enhance collaboration between Research Software Engineers and scientists, ensuring continuous communication and improved outcomes.
- **Status meetings:** Mid-sprint status meetings are now in place to assess progress and tackle any emerging challenges at the one-third mark of the sprint duration.
- **Follow-up evaluations:** A follow-up evaluation, conducted one year after each sprint, will assess the impact and document results, informing future funding opportunities and process improvements.

Resource pooling and collaboration

- **Training and communication initiatives:** Several training sessions on topics like YAC, ComIn, GPU programming, and software engineering have been organized to support our community. Additionally, we are strengthening our communication efforts through newsletters, a dedicated website, and our GitLab platform to promote open development.
- **Encouraging community contributions:** Anja emphasized the importance of showcasing all models, software, and code on our GitLab—even if they aren't selected as core, extended core, or infrastructure components—to foster broader knowledge exchange and collaboration.

Developing our capability and component integration

- **natESM system components:** Our current system is structured into core components (atmosphere, ocean), infrastructure tools (YAC and ComIn), extended core components (e.g., ICON-ART), impact components, and optional components.
- **Focus on impact components:** Noticing a gap in impact components, Anja called for discussions on whether our community should further develop these to better integrate specialized models—such as flood-risk models, biodiversity models, or health-impact models.

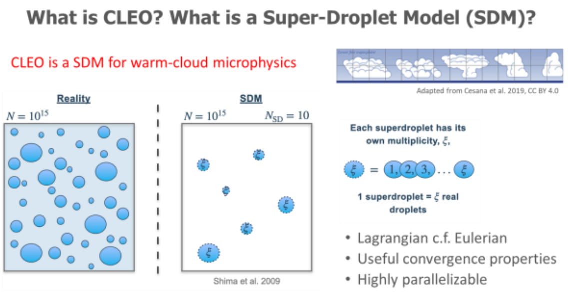
Looking ahead: The natESM2 proposal

- **Future goals and machine-learning integration:** The recent natESM2 proposal aims to align the system more closely with scientific applications by developing well-defined model configurations—like ensemble or air quality simulations—and incorporating machine learning techniques to enhance usability, downscaling, and extrapolation of Earth system data.
- **Community collaboration is key:** Anja highlighted that our achievements to date are the product of voluntary, collaborative efforts. She encouraged more institutions and individuals to engage actively—whether by applying for sprint checks, joining working groups, or sharing innovative ideas—to further drive advancements in Earth system modeling.

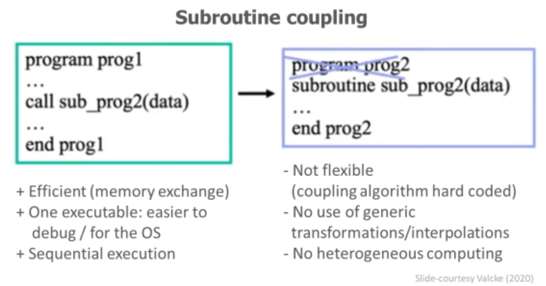
In summary, Anja underscored the value of streamlined processes, enhanced training and communication, strategic framework development, and a forward-looking approach with the integration of machine learning.

Sprint presentations

CLEO Sprint: Clara Bayley from MPI-M presented the results of her sprint that aimed to couple CLEO to ICON using the YAC coupler, marking a key step toward improved interoperability between model components.



modLSMcoup Sprint: Stefan Poll from FZJ presented the results of his sprint, in which the team wanted to develop a proof-of-concept for modular coupling of the land surface and to implement the YAC coupler within the ICON-eCLM framework.



IQ Sprint: Julia Nabel from MPI-BGC presented the results of her sprint that focused on the gradual porting of the IQ code to GPUs. IQ stands for QUINCY in ICON-Land.

	ICON-A		ICON-Land Standalone driver	
	JSBACH	QUINCY	JSBACH	QUINCY (biogeophysics)
Initial status	CPU & GPU		CPU only	CPU only
Porting the ICON-Land standalone driver			CPU & GPU	
Preliminary port of QUINCY (biogeophysics)				CPU & GPU

Following Anja's introduction, we moved straight into the presentations of last year's sprint results. Originally planned as a series of four talks, unforeseen circumstances led to a slight adjustment in the program. Due to illness, only three presentations could be held, but they still provided valuable insights into the progress and achievements of the sprint projects. Despite this change, the discussions that followed were lively and highlighted the impact of collaborative sprint efforts within the community.

After an active afternoon filled with a keynote talk and breakout groups, we gathered for the Icebreaker event, followed by a shared dinner at the conference restaurant. The evening provided a great opportunity to (re)connect, exchange ideas, and get to know new faces in the community—setting the stage for further discussions and collaborations in the days to come.

Engaging conversations and networking over dinner

To the astonishment of many, some individuals—clearly guided by an unwavering sense of virtue—chose to indulge in a wholesome carrot salad for dessert, steadfastly resisting the far more decadent allure of chocolate-caramel pudding.

How AI is revolutionizing climate modeling

Nikolay Koldunov (AWI)

Our first keynote speaker, Nikolay Koldunov from AWI, showcased recent advances in AI and machine learning for weather and climate modeling, highlighting several key developments:

- **AI-Driven weather prediction:** Institutions such as ECMWF, NVIDIA, and Google are developing AI models that forecast weather days ahead. Early comparisons indicate these models achieve forecast skills equivalent to decades of improvements in traditional numerical methods, despite ongoing challenges like field consistency and diffusion.
- **Climate emulators and long-term projections:** Lightweight climate emulators (e.g., Lucy) operate at coarse resolutions and reproduce key climatological features with minimal training data. More advanced coupled AI models—including those incorporating ocean surface temperature—can simulate stable 1000-year rollouts, capturing historical patterns and climate responses to CO_2 variations, though some limitations persist for extreme scenarios.
- **Data processing and communication:** Emerging tools now allow users to interact with extensive climate and observational datasets in natural language. For instance, platforms like Pangaea GPT enable chatbots to generate, execute, and refine code on the fly, simplifying tasks such as species identification or expedition data processing.
- **Public engagement and decision support:** AI systems are being designed to integrate climate data, high-resolution model outputs, and local context (like elevation and demographics) to deliver actionable insights for planning—whether for building sites or wind farms. These interactive, multilingual tools make climate data accessible to non-experts.
- **Overall vision and future prospects:** Focusing on full emulators, data processing, and communication, Nikolay emphasized that despite challenges (like ensuring physical consistency), rapid AI advancements are opening new avenues in weather and climate modeling. He also encouraged engagement with projects like Terra DT, highlighting collaboration and career opportunities in this rapidly evolving field.

With his presentation, Nikolay illustrated how AI techniques can not only enhance weather prediction and climate simulation, but also making complex data more accessible and actionable for scientists, stakeholders, and the public.



YAC, ComIn and beyond – what do you need from natESM?

Florian Prill (DWD), Moritz Hanke (DKRZ), Kerstin Hartung (DLR), Bastian Kern (DLR), Nils Dreier (DKRZ)

The conversations in this breakout group centered on reviewing and comparing the different interface approaches used in ICON and discussing their advantages, limitations, and future directions. The main messages include:

Overview of interfaces: The discussion covered three primary mechanisms—ComIn (a community interface that enables plugin functionality using an embedded Python interpreter), YAC (ICON's coupler library that, among other things, handles model I/O and interpolation between different grids), and the MESSy integration which uses a ComIn plugin to interface with ICON.

For use cases and as guidelines, the group emphasized that:

- **ComIn or plugins** are suitable when you want to use different programming languages (like Python or C++) or when external developers need to work with a stable interface without having to learn the full complexity of ICON.
- **YAC** is particularly useful for handling grid differences, interpolation tasks, and managing modules on heterogeneous hardware.
- A key consideration is whether the code is intended to become a long-term part of ICON or remain an external contribution. If it's external, a plugin-based approach is often recommended because of the complexity of integrating new code into ICON's core.

Challenges and limitations: The group noted that while ComIn reduces interference with the core ICON code, they don't eliminate it entirely. Sometimes, the implementation of a plugin reveals that ICON itself must be modified (for instance, to provide access to temporary variables or to switch off certain parameterization schemes). This inherent limitation means that not all functionality can be externalized purely via ComIn.

Stability and future concerns: There is uncertainty about the long-term stability of these interfaces, especially in light of potential major refactoring projects (like WarmWorld's plans to rewrite parts of the code). The conversation raised concerns about whether the current stable interfaces will remain compatible as the model evolves.

Community and collaboration: The discussion concluded with a suggestion to form a working group to address these interface challenges, exchange experiences (for example, around ComIn plugins), and support the evolution of the interfaces, ensuring that external contributions remain viable despite changes in ICON.

In essence, the conversation was about balancing flexibility (allowing external code and alternative programming approaches) with the need to maintain a stable, integrated core model, while recognizing that some modifications to ICON are inevitable for full functionality.

Land component with focus on “Feedback”

Stefan Kollet (FZJ), Sönke Zaehle (MPI-BGC), Sabine Attinger (UFZ)

The conversation covered two related themes regarding the land component in Earth system modeling, emphasizing both technical challenges and pathways for community collaboration:

Increasing importance of the land component: The discussion began by highlighting that as Earth system models increase in resolution, the detailed representation of land processes (including biogeochemical cycles, disturbances like fires, groundwater dynamics, and human impacts) becomes ever more critical to reduce biases and better simulate global circulation and feedback mechanisms.

Overview of presentations and research areas::

- **Upscaling and biogeochemical cycles:** Presentations addressed the need to upscale plant processes—from molecular to field scales—to improve carbon and nitrogen cycle modeling.
- **Disturbance impacts:** There was a discussion on fire disturbances and their effects on biogeochemistry, illustrating early results and the importance of incorporating such dynamics.
- **Groundwater and hydrology:** A detailed discussion on the continental water equilibrium underlined challenges in modeling groundwater. The participants debated the extent to which detailed hydrological processes (beyond shallow soil moisture and evapotranspiration) affect atmospheric processes and overall model predictions. Key points included the lack of consensus on the necessary complexity of groundwater representation and the challenges due to data scarcity and subsurface heterogeneity.

Community collaboration and future directions: The speakers expressed a strong desire to form working groups and possibly organize dedicated workshops. The goal is to:

- Develop concrete research questions, such as whether including detailed groundwater processes can improve realism or reduce biases.
- Foster collaborations that bridge diverse expertise—ranging from hydrology and biogeochemistry to atmospheric chemistry—to tackle the integration of land processes into ESMs.
- Leverage existing initiatives (like sprint applications) to pilot these improvements in the natESM framework.

Balancing complexity with practicality: There was an ongoing debate on the level of detail required in representing land processes. While more detailed modules might improve scientific realism, they also introduce significant complexity (and potential issues like data limitations). The community is weighing whether to prioritize sensitivity studies and targeted experiments to determine the most effective balance between complexity and model performance.

In summary, the conversation emphasized the need for:

- Better integration of advanced land processes into high-resolution ESMs,
- Focused research efforts (via working groups and workshops) to address both hydrological and biogeochemical challenges,
- A careful evaluation of how added complexity in land process representation can be managed in a sustainable and scientifically productive way.

Toward a new dynamical core for the ICON atmosphere!

Peter Korn (MPI-M)

Our second keynote speaker, Peter Korn from MPI-M, focused on enhancing the dynamical core of the ICON atmosphere model by revisiting and refining its numerical schemes. The main messages included:

Reevaluating foundational algorithms: Peter explained that many seminal papers on dynamical cores, once groundbreaking, are now over a decade old. He argued that numerical methods for shallow water, hydrostatic, and nonhydrostatic models can lose efficiency and robustness over time. In light of new hardware—especially GPUs—and modern computational needs, he called for a thorough reexamination of these methods.

Focus on conservation properties: A key theme was the importance of maintaining discrete conservation properties such as momentum, energy, vorticity, helicity, and angular momentum. Peter demonstrated that proper reconstruction techniques and careful discretization—handling cross products and vector invariants—can preserve these physical symmetries (reflecting ideas from Noether's theorem) and lead to more robust simulations.

Handling nonlinearity and flow regimes: Peter stressed that the model should remain robust across different flow regimes. The numerical formulation must allow a seamless transition between compressible and incompressible, or nonhydrostatic and hydrostatic regimes, without compromising conservation.

Practical and computational considerations: While incorporating advanced reconstruction and conservation techniques might seem to add complexity, Peter argued that these changes are localized (affecting roughly one module) and do not significantly increase computational cost. In fact, models that honor conservation laws tend to have fewer numerical instabilities and require less tuning, as seen in comparisons with the ocean model ICON-O.

Looking ahead—modernization and collaboration: Peter emphasized that merely porting old algorithms to new architectures like GPUs isn't enough. Instead, a rejuvenated platform or communication forum is needed to share, develop, and benchmark improved methods. He also hinted that aligning parameterizations of unresolved physics with the enhanced dynamical core could further boost model performance.

In summary, Peter advocated for a comprehensive reexamination of core numerical algorithms—emphasizing conservation, symmetry, and robust asymptotic behavior—to develop more reliable and efficient dynamical cores for modern high-resolution modeling on advanced computational architectures.

Impact modeling and land feedback

Sabine Attinger (UFZ),
Stefan Kollet (FZJ), Sönke Zaehle (MPI-BGC)

Sabine Attinger explained that many impact modelers—from forestry, crop, and urban sectors—have stepped back, partly because they rely on pre-processed meteorological data rather than raw climate model outputs. She stressed that demonstrating how natESM outputs can directly benefit their work is essential for re-engaging these experts. To achieve this, she proposed several measures:

Proposing a working group on data processing

- Discuss and share methods for bias correction and downscaling.
- Enhance access to existing datasets (using tools like Easy Map) and work toward standardizing processing steps.

Hydrology and land-surface schemes

- Hold a bilateral meeting with Sönke to review hydrology in land-surface models—particularly within the ICON Land framework—and explore potential improvements without forming a full working group.

Improving landscape data accessibility

- Emphasize the importance of high-quality landscape attributes (e.g., topography, soil data) and reliable transfer functions to map these data to model parameters.
- Form a working group to harmonize global, regional, and national datasets and share existing methodologies, so modelers do not have to start from scratch.

In summary, Sabine's group called for structured collaboration through two targeted working groups and bilateral conversations to address data and methodological gaps—efforts that could help reconnect impact modelers with the natESM community.

Atmospheric chemistry and aerosol modeling

Corinna Hoose (KIT),
Ina Tegen (TROPOS), Anja Schmidt (DLR, LMU)

In this group, participants discussed the future direction and unification of aerosol-chemistry modeling within the natESM system. The main points were:

Current status and motivation

- Three major aerosol chemistry models are currently in use within the German community.
- The group questioned whether maintaining these models in parallel is sustainable and resource-efficient.

Vision for the future

- Participants were encouraged to envision what an aerosol chemistry model should achieve in the next 5–10 years.
- A unified approach could enable new scientific inquiries and use cases—such as exploring feedbacks across different time scales (e.g., biogeochemical cycles, climate risks)—that are not currently feasible.

Technical and scientific requirements

- Seamless complexity and consistency: The unified model must maintain consistent physics and chemistry across various scales, ensuring coherence when changing resolutions or parameterizations.
- Modularity and adaptivity: Users should be able to configure and adapt different modules (e.g., switch specific processes on or off) based on simulation needs.

Future steps and coordination: Although many valuable ideas were raised, no definitive decisions were made at this stage. The speakers will serve as contact points to gather feedback, refine use cases, and coordinate follow-up, potentially through a working group or workshop.

Summary and outlook

Jochem Marotzke (MPI-M)

The closing summary by Co-Chair Jochem Marotzke highlighted several overarching themes and actionable points from the workshop. The main takeaways were:

Widening use cases and expanded functionality: Tools like ComIn and YAC are now applied in an ever-growing range of contexts, extending well beyond their original design and reflecting the community's diverse needs.

Interference with host models: Although ComIn has substantially reduced direct interference with the ICON host code, some interference remains. Ongoing dialogue between the community and ICON developers is necessary to determine what level of code modification is acceptable.

Community engagement and impact modeling: The group noted a decline in impact modelers' participation and stressed the importance of initiatives—such as workshops or working groups—to revitalize this segment of the community and demonstrate the value of natESM outputs.

Technical challenges and future directions

- Portability and hardware: A major challenge is ensuring code portability across different hardware (e.g., NVIDIA vs. AMD GPUs), with vendor lock-in a significant concern.
- Programming environment: Choosing programming languages and tools (e.g., potentially moving from Fortran to Python or Julia) will be crucial for attracting new talent.
- Hydrology: There is interest in exploring the explicit representation of groundwater hydrology. While its ultimate significance is still uncertain, it may yield valuable insights.

Driving change with a big use case: A key takeaway was the need for a compelling "big carrot"—a high-impact use case that requires capabilities not yet possible. Such a unifying challenge could drive the community to develop a next-generation aerosol and atmospheric chemistry model.

Encouragement for determined progress: Jochem urged the community to embrace this challenging path with open eyes, recognizing that while it may be thorny and labor-intensive, the long-term benefits—much like those observed in the evolution of ICON—are well worth the effort.

All workshop presentations and slides can be found on **our website**.

Moving forward

In total, the workshop discussions led to **four proposed working groups** and **three focus workshops**, reflecting the community's commitment to tackling key challenges collaboratively:

4 Working groups: ComIn-plugin exchange (Prill), Groundwater in land feedbacks (Kollet), One-way coupling of climate and impact models (Attinger), and landscape attributes for land-surface modules (Attinger).

3 Focus workshops: Groundwater in land feedbacks (Kollet), Hydrology in land biogeochemistry (Zaehle), and ocean biogeochemistry (Hauck).

We look forward to addressing these challenges together in the coming months!

📢 ESiWACE3 Call for Proposals

The ESiWACE3 project is **actively seeking applications** for software support in climate and weather modeling. Don't miss this opportunity! Apply here: **ESiWACE Call**

💻 Global Climate Modelling Hackathon

Join the Hamburg Node!

Taking place **May 12-16, 2025**, this global hackathon pushes the boundaries of **cutting-edge climate modeling**. Open to all—including experienced researchers looking to refresh their skills!

🎓 hpc4climate Summer School

From **July 28 – August 7, 2025**, WarmWorld & ESiWACE invite **Master's and early PhD students** in CS, Data Science, Math, and Climate Science to **explore scientific computing in HPC**.