natESM white paper: Uniting Earth System Modeling for a Sustainable Future

Iris Ehlert[1], Sabine Attinger[2], Hendryk Bockelmann[1], Georg Feulner[3], Corinna Hoose[4], Thomas Jung[5], Stefan Kollet[6], Roland Potthast[7], Kira Rehfeld[8], Hauke Schmidt[9], Ina Tegen[10], Sönke Zaehle[11], Jochem Marotzke[9], Anja Schmidt[12,13]

[1] Deutsches Klimarechenzentrum GmbH, Hamburg, Germany
[2] Helmholtz-Zentrum für Umweltforschung GmbH, Leipzig, Germany
[3] Potsdam Institute for Climate Impact Research, Potsdam, Germany
[4] Karlsruhe Institute of Technology, Karlsruhe, Germany
[5] Alfred-Wegener-Institut, Bremerhaven, Germany
[6] Forschungszentrum Jülich GmbH, Jülich, Germany
[7] Deutscher Wetterdienst, Offenbach, Germany
[8] Universität Tübingen, Tübingen, Germany
[9] Max Planck Institute for Meteorology, Hamburg, Germany
[10] Leibniz Institute for Tropospheric Research, Leipzig, Germany
[11] Max Planck Institute for Biogeochemistry, Jena, Germany
[12] Deutsches Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen-Wessling, Germany
[13] Meteorological Institute, Ludwig Maximilian University of Munich, Munich, Germany

Version as of 26 January 2024

Welcome to natESM, the initiative dedicated to transforming Earth system modeling in Germany by fostering collaboration among research institutions, public authorities, and universities. By pooling national resources, we aim to accelerate and share scientific and technological developments. The advent of exascale high-performance computing in the mid-2020s presents significant challenges, necessitating innovative model and software development due to heterogeneous hardware architectures. Our vision is to establish a world-leading, ready-to-use Earth system modeling capability, enabling the German ESM community to leverage progress in high-performance computing. This document details the natESM strategy, guiding principles, and the path toward realizing this ambitious vision.

Contents

The natESM project is supported by the PT-DLR and funded through the Federal Ministry of Education and Research (BMBF) under grant agreement No. 01LK2107A.

1 Vision and mission

At the heart of natESM lies a compelling vision: to create a world-leading, multiscale, seamless, ready-to-use Earth system modeling capability that serves research, operational applications, training, and education. natESM aims to harness German expertise while fostering international collaboration without creating dependencies, and embracing transparent governance.

Our mission is equally ambitious: to cultivate a community where resources and expertise are shared, paving the way for the next generation of Earth system modeling capabilities. Our community includes anyone with an interest in natESM, regardless of their level of involvement, whether they are core contributors or simply share a passion for Earth system modeling and appreciate the user-friendly nature of the future natESM capability.

2 natESM values: the foundation of our strategy

We connect our vision with our mission through the following values that underpin our strategy.

1. **Reciprocity**
   We believe in the power of collaboration, where knowledge, insights, and resources are exchanged to foster innovation and advancement.
2. **Openness**
   Open science is our ethos. Open access, open data, open-source principles, and open development are fundamental to natESM, ensuring transparency and accessibility.
3. **Scientific diversity**
   We embrace diverse perspectives, competencies, and experiences, recognizing the strength they bring to our community.
4. **Congruence of institutional and collective goals**
   We align our individual and collective objectives to create a harmonious and effective collaborative environment.
5. **Agility: adapting to change**
   Our strategy is designed to be agile, acknowledging the inevitability of change in the ever-evolving world of Earth system modeling and high-performance computing. This ensures our ability to adapt, evolve, and thrive in this dynamic environment.

3 Key components shaping natESM's evolutionary journey

natESM's strategy fosters a coherent and effective development path, ensuring that contributions positively impact the system as a whole. In this section, we delve into the crucial elements.

3.1 Collaborative Earth system modeling components: aligning with core components

natESM comprises four types of Earth system modeling components, designed to adhere to levels of well-defined interfaces and rigorous technical criteria. For each individual model or module, the institution developing the respective component is responsible for the maintenance and further development of its component.

Core components
These form the foundational elements vital for Earth system modeling, acting as prerequisites for most other models and modules. Our current core components encompass models for the atmosphere (ICON-A) and the ocean (ICON-O and FESOM), with ongoing discussions regarding the selection of suitable models for the land, atmospheric chemistry, and land-ice components. It is the responsibility of natESM to seamlessly integrate these core components into the natESM system.

Optional components
These components, although valuable, are typically not required by the majority of users. They remain optional, providing flexibility for specific use cases. natESM supports their integration into the modeling system.

Impact components
Primarily facilitating one-way interactions within the system, natESM supports the integration of these components into the modeling system.

Infrastructure components
Covering various aspects like input/output, coupling, and interfaces, these components enable the interaction with the core and optional components, ensuring the seamless and convenient functioning of the system. Our current infrastructure components include a coupler (YAC) , suitable for all models, and the application programming interface ComIn for accessing data structures and variables in ICON. It is natESM's responsibility to oversee the integration of these infrastructure components, and potentially others, into the envisioned natESM capability.

You can find a comprehensive overview of the components on our GitLab. Guidance on integrating new models or components into the natESM system is vital, with a focus on alignment with our core components without imposing undue restrictions. Since natESM welcomes contributions that meet specified criteria, it is crucial to consider how these additions harmonize with or enhance the capabilities of our core components.

3.2 Criteria for acceptance

The criteria for acceptance into the natESM capability include:

1. **Alignment with the German community**
   Models should ideally align with the national research objectives and priorities of the German Earth system modeling community.
2. **A long-term institutional commitment**
   Institutions must express their commitment to the continuous development and support of their models. For university contributions, natESM offers an adoption mechanism (see section 3.3).

3. **Technical criteria**
   Models must conform to agreed-upon technical criteria (see Appendix A1 for details), ensuring compatibility and reliability within the natESM system.

These criteria aim to ensure that accepted models or modules enhance natESM's capability and align with the strengths of our core components, thereby maintaining a coherent and effective strategy that advances Earth system modeling in a unified and productive manner.

3.3 Inclusion of university contributions

Universities often encounter challenges when committing to long-term Earth system model development due to staffing and funding constraints, especially within short-term research projects or Ph.D. programs. To address this, natESM offers an adoption mechanism, allowing institutions within our community with a professional interest in a university-developed module or model to adopt it.

This adoption process ensures compatibility with natESM's technical criteria, offers long-term commitment, and nurtures continued development. The terms and conditions of continued university involvement and how they receive credit after adoption are negotiated on a case-by-case basis between the university and the adopter.

Embracing this concept ensures that university contributions are not only welcomed but also sustained and nurtured within our collaborative ecosystem, promoting inclusivity, and recognizing our community's diverse strengths and constraints.

3.4 Sprints: fueling natESM's progress

Sprints are the lifeblood of natESM, serving as essential catalysts for innovation and technical progress. They are integral to our mission of developing a world-leading Earth system modeling capability. Sprints empower institutions to refine their models and modules, ensuring they meet the rigorous technical criteria necessary for seamless integration into the system. During these intensive working periods, experts collaborate closely to improve the performance, scalability, and functionality of our Earth system models. By providing expert guidance and support, sprints enable us to collectively enhance the performance, scalability, and functionality of our Earth system models. Through these collaborative efforts, we not only drive scientific advancements but also strengthen the foundation of natESM as a cutting-edge research initiative (see Appendix A2 for details).

3.5 Community engagement and adaptive communication

In the natESM initiative, community engagement and effective communication are intrinsically linked. Our community encompasses anyone interested in natESM, irrespective of their level of involvement—whether they are core contributors or simply passionate about Earth system modeling and the user-friendly aspect of the future natESM capability. We actively engage with our community through various avenues, providing a platform for sharing insights, exchanging knowledge, and collaboratively shaping our project's direction.

1. Working groups

Working groups play a crucial role in advancing natESM's objectives by focusing on specific technical and scientific goals. They enable community members to actively contribute to the project's development and facilitate collaboration among members with diverse structures and resources, ensuring their ideas and professional expertise are readily available to the entire community (see Appendix A3 for details).

2. Online communication

Simultaneously, our adaptive communication strategy is instrumental in reaching our diverse community. We leverage multiple media channels, including email, a quarterly newsletter, a dedicated Mattermost channel, and a website, ensuring smooth and inclusive information flow. Additionally, we offer a GitLab platform where community members can showcase valuable models or modules, encouraging active participation and feedback.

3. Workshops & trainings

Our flagship event, the **annual community workshop**, serves as a nexus for the entire natESM community. During this gathering, we consolidate insights from sprints, working groups, and smaller workshops conducted throughout the year, for collective discussions on upcoming challenges and advancements. These discussions help shape the direction of natESM, ensuring that the entire community is involved in the project's ongoing evolution.

In addition to the broader annual workshop, we organize smaller, **focused workshops** as topics emerge. These workshops are not pre-planned a year ahead; instead, they unfold organically based on community needs. Whether sparked by individual ideas, steering group initiatives, or community member requests, these workshops provide a flexible space for in-depth discussions on specific components and emerging challenges within the Earth system modeling system.

Complementing our workshop series, we conduct **technical training** sessions designed to empower our community with the knowledge and skills needed to navigate the natESM system effectively. These sessions cover diverse topics, ranging from system usage and coupling strategies to in-depth explorations of interfaces. Our goal is to equip community members with the tools they need to harness the full potential of natESM for their scientific inquiries.

Through this holistic approach, we foster collaboration, share expertise, and empower our community to actively contribute to the advancement of Earth system modeling within natESM.

4 Conclusion

natESM is more than a project; it is a community united by a shared vision, mission, and values. Our strategy reflects our commitment to innovation, collaboration, and openness in Earth system modeling. We invite research institutions, public authorities, and universities to join us in this transformative journey as we work together to create a world-leading Earth system modeling capability.

Appendix

Contents

A 1 Technical criteria

The technical criteria outlined for becoming a part of the natESM system provide a comprehensive framework for ensuring the efficiency, adaptability, and accessibility of the Earth system modeling components. Here's a breakdown of these criteria:

1. **Well-defined interfaces between Earth System Components**
   Components within the natESM system should have clear and well-documented interfaces to facilitate seamless communication and data exchange between different parts of the model.
2. **Allows simulations from global to local scales**
   The modeling system should support simulations across various scales, ranging from global-scale climate modeling to more localized and region-specific simulations.
3. **Exascale-ready**
   The system should be prepared to harness the computational power of exascale supercomputers, ensuring it can handle the massive computational demands of high-resolution simulations.
4. **Scalable workflows**
   Scalability is essential to adapt to various computational resources and research needs. Workflows should be designed to efficiently use available resources while maintaining performance.
5. **Portability**
   The modeling system should be portable across different computing platforms and environments, making it accessible to a wide range of users and institutions.
6. **Modularity**
   A modular design allows for flexibility and the ability to swap out or update specific components without affecting the entire system's functionality.

7. **Data assimilation capacity**
   For those interested in making forecasts, the system should be capable of assimilating observational data into simulations to improve the accuracy and realism of model outputs.
8. **Diagnostic capacity**
   Robust diagnostic tools should be integrated into the system, enabling users to analyze and interpret simulation results effectively.
9. **User-friendly and well-documented**
   A user-friendly interface and comprehensive documentation are essential to ensure that researchers and scientists can easily access and understand the system.
10. **Traceability, reproducibility, and version control**
    Traceability ensures that simulation results can be traced back to specific input conditions and configurations. Reproducibility is crucial for scientific rigor, allowing others to replicate research findings. Version control helps manage updates and changes to the system.
11. **Standardization**
    Standardization of processes and data formats promotes consistency and interoperability across different components and institutions.
12. **License of Useful Open-Source Type**
    The software and components used in the natESM system should be released under open-source licenses that allow for collaboration, modification, and redistribution while protecting intellectual property rights.

These technical criteria collectively emphasize the importance of a flexible, high-performance, and user-friendly Earth system modeling capability. By adhering to these criteria, natESM aims to create a robust and adaptable platform that can advance Earth system research and facilitate collaboration among a diverse group of institutions and researchers.

A 2 Exploring sprint opportunities with natESM

natESM maintains an open call for proposals, enabling model-development groups across Germany to become part of our Earth-system research community.

Our sprints, focused on technical objectives and tethered to natESM resources, provide a flexible program tailored to your research goals and timelines. This collaborative journey spans up to six months, fostering in-depth partnerships between you and our Research Software Engineers (RSEs).

A 2.1 The sprint process: from sprint check to report

The sprint process within natESM encompasses several stages, ensuring that each sprint effectively contributes to our Earth system modeling objectives. Here is an overview of the sprint journey that you can initiate at any time:

1. **Initiate the process through a sprint-check request**
   Prospective applicants are required to undergo a sprint check, serving as an accessible entry point for guidance before submitting a full sprint application. If you are contemplating a sprint, send a brief description of the challenge you

intend to address, along with access to the code for our RSEs and details about your experimental setup to support-request@nat-esm.de. This essential, low-threshold step is designed to assist you in initiating your sprint journey promptly and establish a connection with the RSEs who will guide you throughout this process.

If the software or model is suitable for improvement within a sprint, the RSEs will provide precise advice on what should be included in a full proposal and how it should be structured before submission.

In cases where a sprint is not currently feasible due to outdated code, the RSEs will provide you with detailed information on how to improve the code. Once you have incorporated the RSEs' suggestions, you can contact us again at any time with a sprint-check request.

Typically, our RSEs conduct sprint checks concurrently with their dedicated work on an active sprint. The timeframe for completing a sprint check ranges from as short as two weeks to a maximum of two months.

Please note: A successful sprint check with a recommendation to submit a full proposal does not imply that the proposal will be accepted!

2. **Application submission**
   Researchers submit comprehensive sprint applications, considering advice from the sprint check, detailing proposed goals, scope, and expected outcomes to support-request@nat-esm.de.
3. **Expert review**
   The RSEs, the coordinator, and steering-group members conduct a thorough review of full sprint applications over approximately six weeks.
4. **Notification**
   Applicants are notified of the outcome by email. If accepted, a KickOff meeting is set up to initiate collaboration.
5. **KickOff meeting**
   The official start of the sprint, where the appointed RSE is introduced. This meeting establishes initial goals, communication frameworks, and team dynamics.
6. **Progress check**
   For sprints with a duration of 6 months, about two months into the sprint, a status meeting is held to assess progress and address any challenges. This check ensures that the sprint stays on track and aligns with its goals.
7. **Sprint report**
   After the sprint, applicants collaborate with RSEs to prepare a detailed report, highlighting achievements, outcomes, and lessons learned. Published on the natESM website within three months, the report provides valuable insights for the natESM community.

This structured sprint process fosters effective collaboration, accountability, and transparency throughout the sprint's lifecycle, ensuring that each sprint contributes meaningfully to the advancement of Earth system modeling within the natESM capability.

A 3 Establishment and dissolution of a working group

Working groups provide a mechanism for community members to actively contribute to the natESM development and scientific pursuits by focusing on specific technical or scientific objectives.

Compared to sprints, which focus primarily on technical tasks and do not have a specific scientific objective, working groups have a broader scope, encompassing both scientific and technical objectives. Working groups should be active for no longer than 12 months.

Working groups can be proposed by anyone from the community who volunteers to take the lead and to serve as the contact point for others interested in attending the group.

1. A working group always has a clear goal that helps advance natESM.
2. A working group should not last longer than one year.
3. Achieving a technical development goal enables the simulation needed to pursue a scientific goal.
4. A technical development goal should be confined enough to permit rapid completion.
5. Additional technical development goals should be tackled in a future working group.
6. A proposal for a working group can be made through a brief write-up to the process coordinator summarizing the working-group goals, as well as a schedule.
7. A working group is completed through a brief retrospective describing outcomes and lessons learned.
8. Working-group proposals and retrospectives are made available on the natESM website.

A 4 The natESM steering group

The natESM steering group, comprised of the community members listed at the beginning of this document, plays a pivotal role in shaping the strategic direction, governance, and overall natESM capability. This dedicated team is responsible for defining core components and guiding the establishment of working groups, ensuring alignment with natESM's strategy. Additionally, the members of the steering group represent the core components of natESM and bear institutional responsibilities, reinforcing the commitment to the effective and coordinated advancement of the natESM capability.

A 5 Institutions already active in natESM

Several institutions are already actively contributing their expertise, resources, and models to collaboratively advance Earth system modeling within the natESM capability. To explore the list of these committed members, visit our website.

If your institution is interested in supporting natESM and would like to be featured on this page as well, please don't hesitate to send a brief email to info@nat-esm.de. We welcome all institutions with an interest in contributing to our shared vision.