https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


# A case for open communication of bugs in climate models, made with ICON version 2024.01

Ulrike Proske[1], Nils Brüggemann[2], Jan P. Gärtner[3], Oliver Gutjahr[2], Helmuth Haak[2], Dian Putrasahan[2], and Karl-Hermann Wieners[2]

[1]Hydrology and Environmental Hydraulics Group, Wageningen University, Wageningen, the Netherlands
[2]Max Planck Institute for Meteorology, Hamburg, Germany
[3]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

**Correspondence:** Ulrike Proske (ulrike.proske@wur.nl) and Oliver Gutjahr (oliver.gutjahr@mpimet.mpg.de)

**Abstract.** Climate models are not just numerical representations of scientific knowledge, they are also human-written software programs. As such, they contain coding mistakes, which may look mundane, but can affect the results of interconnected and complex models in unforeseen ways. These bugs are underacknowledged in the climate science community.

We describe a sea ice bug in the coupled atmosphere-ocean-sea ice model ICON and its history. The bug was caused by a logical flag that was set incorrectly, such that the ocean did not experience friction from sea ice and thus the surface velocity did not slow down, especially in the presence of ocean eddies. While describing the bug and its effects, we also give an example of visual and concise bug communication. In addition, we conceptualize this bug as representing a novel species of resolution-dependent bugs. These are long-standing bugs that are discovered during the transition to high-resolution climate models due to features that are resolved at the kilometer scale. This case study serves to illustrate the value of open documentation of bugs in climate models and to encourage our community to adopt a similar approach.

## 1 Bugs in Climate Models

Climate models are not only digital manifestations of physical laws, they are also human-written software programs. Like any software, climate models contain errors and mistakes made by modelers, called bugs. These are to be expected, because human activity is error-prone (Newman, 2016). Bugs have been shown to be common in commercially released software (Hatton, 1997; Soergel, 2015), and to be a problem for scientific software in general (Miller, 2006; Hatton, 2007; Barnes and Jones, 2011; Soergel, 2015) and for environmental and climate models in particular (Luther et al., 1988; Menard et al., 2021).

There is a large body of literature on scientific software and its relationship to software engineering practices (see e.g. Segal (2004); Wilson (2006); Carver et al. (2007); Kendall et al. (2008); Basili et al. (2008); Kelly and Sanders (2008); Segal (2009); Kelly (2015); Paine and Lee (2017); Storer (2017); Tucker et al. (2022); Depaz (2023)). In a pioneering study, Hatton (1997) looked for bugs in several commercially used scientific software programs, such as one for seismic data processing in the Earth Science industry. He found that all of the investigated software he examined contained bugs, some so serious that they rendered the results of the program useless. Various types of bugs have been identified in the literature, ranging from compiler (Rahman et al., 2023) and language faults (Hatton, 1997) to numerical bugs (Di Franco et al., 2017) and to scalability bugs (Calotoiu

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


et al., 2013). Di Franco et al. (2017) found that correctness bugs are the most common numerical bugs. Examples include typographical errors or using a function outside its scope.

An important distinction for climate models is that between acknowledged and unacknowledged errors. Acknowledged errors "are unavoidable or intentionally introduced to make a problem tractable". Examples of this type are physical parameterizations, which are not considered bugs in common understanding but are rather attempts to represent sub-grid scale processes. Unacknowledged errors, however, "result from blunders or mistakes" (Hook and Kelly, 2009; Pipitone and Easterbrook, 2012). What makes it difficult to find unacknowledged errors is the complexity of the model (Stensrud, 2007; Chowdhury and Zulkernine, 2010; Hajek and Wolinsky, 2012; Pipitone and Easterbrook, 2012; Mendez et al., 2014; Larsen et al., 2016). Moreover, the results of these models are not known at the start, meaning that there is rarely a known truth against which to test the results. This dilemma is known as the oracle problem (Hook, 2009; Hook and Kelly, 2009). At the same time, acknowledged errors are difficult to bound and enlarge the tolerance in output testing, which can hide unacknowledged errors. Hook (2009) calls this the tolerance problem. In addition, noticing, examining, and fixing bugs requires domain knowledge (Di Franco et al., 2017), but climate scientists learn programming by doing, and rigorous software training is often lacking in traditional scientific education (Sundberg, 2009; Merali, 2010; Barnes and Jones, 2011).

Focusing on climate models, Easterbrook and Johns (2009) approximated a defect density of 2 per 1000 changed source lines of code that pass through testing and review for each release of the Met Office's Unified Model. In a follow-up study, Pipitone and Easterbrook (2012) analysed defect reports and fixes in several global climate models. They found that the models had "very low defect densities compared to well-known, similarly sized open source projects." They attribute this low defect density to the modelers' domain expertise, their rigorous development process, and caution, but also to their disregard for high technical standards as long as the code produces correct results. Pipitone and Easterbrook (2012) highlight the need for more in-depth studies of the defects that climate modelers face and how they are found.

Climate modelers themselves usually do not publish what does not work. Bugs may be used and published as sensitivity experiments or documented on the side as part of a new model version (see for example Tallapragada et al. (2014); Jöckel et al. (2016); Graff et al. (2019); Bethke et al. (2021); Mauritsen et al. (2022); Shaw et al. (2022); McGraw et al. (2023)), but are rarely the subject of a study. Bugs are described only to the extent necessary to explain the characteristics and caveats of particular simulations. A notable exception is Kahre et al. (2023), who devoted an entire paper solely to documenting a bug they found in the radiation code of the Legacy Mars Global Climate Model, and detailing the effects on the results and conclusions of an earlier paper. Similarly, Shaw et al. (2022) and Zhu et al. (2022) openly discuss an erroneous cloud ice number concentration limiter, whose removal significantly lowered the climate sensitivity of the Community Atmosphere Model.

We expand on the above work with a case study detailing a specific bug we found in the coupled atmosphere-ocean-sea ice Sapphire configuration of ICON. Our contribution is three-fold. First, we document a sea ice-ocean interaction bug in ICON. This documentation is useful for researchers using ICON, as this bug has been present since March 2015 and is therefore present in numerous simulations and articles. We aim to document the scientific and epistemic richness contained in the history of the bug, making the debugging process explicit (see Sec. 3.2 and 4). Second, in Sec. 3 we give an example of how to publish bug reports and fixes. Case studies of bugs are rare in the climate modeling literature, but we believe that understanding our

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


models requires understanding their code errors. Our goal is to provide an example of visual and concise documentation that will aid internal communication and allow researchers in different modeling projects to learn from our mistakes. We hope to demonstrate the value of open reporting and encourage our community to report bugs. Third, we conceptualize our bug of interest as part of a new category of resolution-dependent climate model bugs (see Sec. 5). It is one of several bugs discovered by the ICON Sapphire development team while building and running the model at storm-resolving resolution. We noticed that long-standing bugs were only discovered because of features resolved at kilometer scale. We discuss the implications, as these types of bugs will become increasingly important as many modeling efforts move to higher resolutions.

## 2 Sea Ice and Ocean Model

Sea ice plays an important role in the climate of the Earth, regulating the exchange of mass and energy between the polar oceans and the atmosphere and redistributing salt in the ocean. Due to its high albedo, sea ice cools the polar oceans and the atmosphere by reflecting much of the incoming short-wave solar radiation. When the sea ice melts completely, the underlying darker ocean is warmed by absorbing shortwave radiation due to its lower albedo.

Sea ice grows thermodynamically at the ice-ocean interface as congelation ice due to the conductive heat flux from the ocean to the atmosphere through the ice. The heat loss of the ocean is thus balanced by freezing at the ice bottom. New ice can also form in the open water when the water column is at freezing temperature, such as in polynyas or leads. Sea ice can also increase in thickness mechanically as ice floes slide over each other, called rafting for thin ice or ridging for thick ice. Sea ice thickness is therefore an irregular field with large spatial variability, ranging from thin ice in leads and polynyas, where most of the new ice forms, to several meters in hummocks and pressure ridges.

The drift of sea ice is governed by the stresses exerted by the wind and the ocean on the ice mass, as well as by the tilt of the sea surface height and the deflection by the Coriolis force. If only these forces act on the sea ice, it is in free drift. In regions where the sea ice is influenced by boundaries, such as land, or regions of immobile sea ice, such as landfast ice, the sea ice is no longer in free drift and may deform. Areas of frequent deformation are therefore found in shear zones. Deformation is controlled by the thickness and compactness of the ice. Typically, wind stress is the dominant force in the sea ice momentum, while ocean stress slows the ice by friction, which in turn drifts at an angle to the wind direction.

In ICON's ocean and sea ice model, a continuum assumption is made such that a grid cell contains multiple floes of different sizes, and the sea ice velocity is calculated based on the average ice thickness within a grid cell. The coverage of an ocean grid cell by sea ice is called the sea ice concentration and ranges from 0 to 1 (or 0 to 100 %). The stresses from the wind and ocean are typically formulated as quadratic drag laws (see Sec. 3), and the deformation of sea ice is formulated as a rheology (see Sec.2.1 for details on the sea ice component of ICON).

### 2.1 Model configuration and experiment

We analysed a coupled simulation produced with ICON, in the Sapphire kilometer-scale configuration (Mauritsen et al., 2022; Hohenegger et al., 2023a). The philosophy of ICON-Sapphire is to use as few parameterizations as possible, for the atmosphere,

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


land, and the ocean/sea ice components. The simulation was conducted within the nextGEMs project (nextGEMS, 2024) and ran for 30 years, from January 1st 2020 to January 1st 2050. The atmosphere component including the land module JSBACH (Reick et al., 2021) had a resolution of about 10 km (R2B8) and 90 vertical levels. The ocean/sea ice component (ICON-O, Korn (2017); Korn et al. (2022)) used a horizontal cell size of about 5 km in the linear dimension (R2B9) and 72 vertical levels. The atmosphere and ocean components were coupled every 10 minutes. The sea ice dynamics were adapted from FESIM (Danilov et al., 2015) with an elastic-viscous-plastic (EVP) rheology (Hunke and Dukowicz, 1997) and a replacement pressure (Kreyscher et al., 2000; Kimmritz et al., 2017), with an elliptic yield curve and ice strength according to Hibler (1979). Sea ice thermodynamics use a zero-layer model (Semtner, 1976), with one ice-thickness class, allowing for fractional ice cover of a grid cell. Key features of the simulation include explicit resolution of deep convection in the atmosphere, mesoscale eddies in the ocean, and linear kinematic features in the sea ice.

The low-resolution model version employed for comparison used the extended prediction and projection ICON configuration (Müller et al., 2024). The configuration was designed to be comparable to the high-resolution experiment, with a few noteworthy differences. The coarser atmospheric grid had a horizontal resolution of 80 km and 130 vertical levels. In contrast to the high resolution configuration, the numerical weather prediction (NWP) physics package with a different radiation code, turbulent mixing, and the usual set of active parameterizations (e.g. shallow and deep convection) was used. Note that while both atmospheric model versions were developed within the ICON consortium, this NWP package comes from a different line of development (Zängl et al., 2015). The ocean component used a grid with 20 km horizontal resolution and the same 72 vertical levels as the high resolution setup. Besides the coarser horizontal resolution and an active mesoscale eddy parameterization, other ocean parameters (e.g. sea ice, vertical mixing) were kept as close as possible to the high resolution setup. As in the high resolution, the ocean was initialised from ORAS5 ocean reanalysis (Zuo et al., 2019) for January 1st 2010. The high-resolution ocean model ran for 10 years as spin-up prior to coupling on January 1st 2020, whereas the low-resolution model was coupled from the start. It then ran transient for 30 years, using similar forcing (greenhouse gases, aerosol) as the high-resolution configuration.

## 3 Results

The sea ice hole appeared in a high-resolution coupled simulation in which the atmospheric model was coupled interactively with the ocean-sea ice model (see Sec.2.1), as visualized in Fig. 1a. As explained in section 1, we believe that open communication of bugs is beneficial for the scientific community. We propose Fig. 1 as a template for bug communication. It summarizes all aspects important to understand a given bug: in which configuration the bug occurs, the effects or visual manifestation of the bug, how the bug is included in the code, how it was fixed, and the results after the bug fix. We elaborate these aspects in detail below. In principle, the graphic should be sufficient to give a model user an overview of the bug and whether the bug may have an influence on their research using the model results. Further ideas for and caveats of the graphic are discussed in Sec. 5.1.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


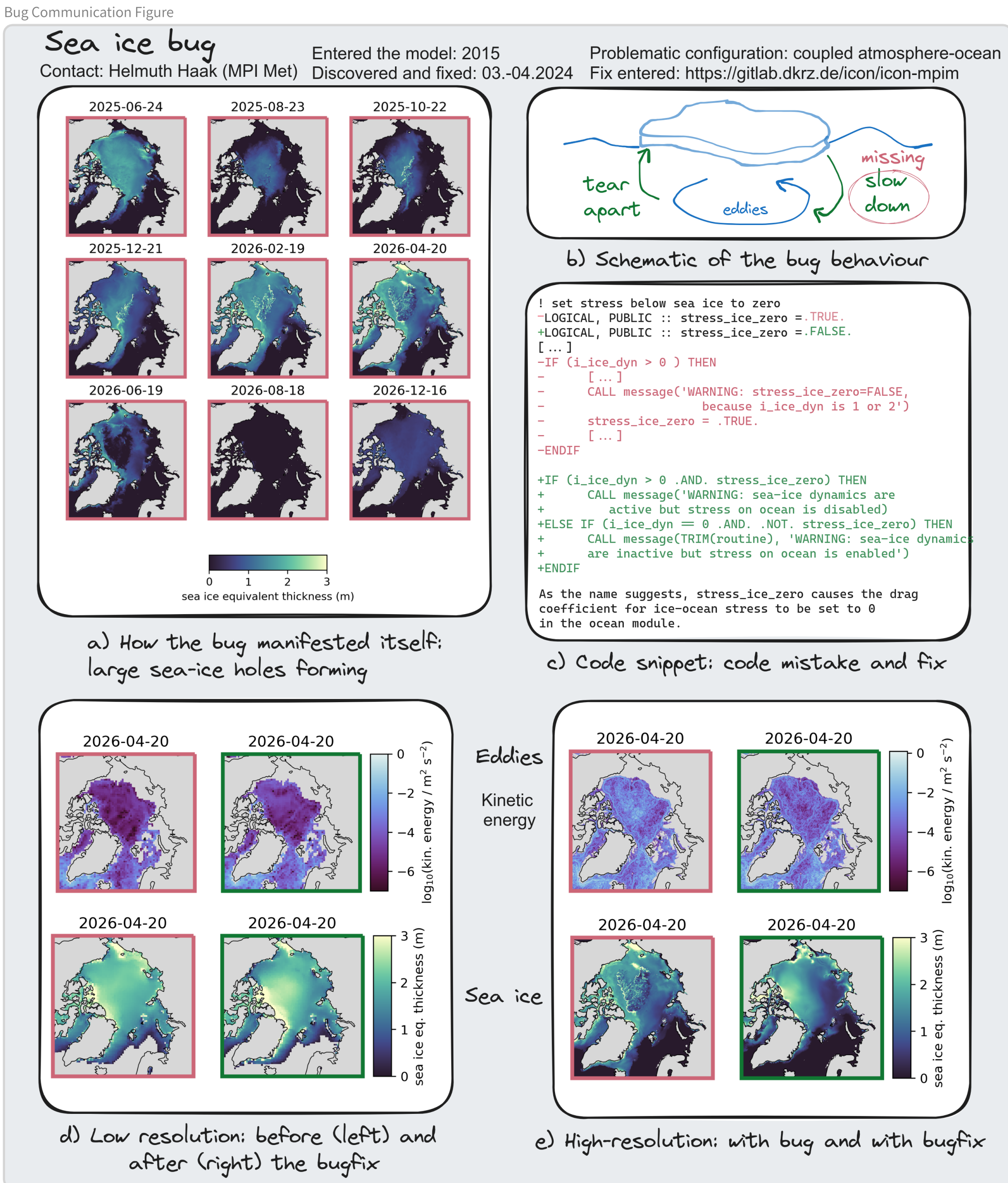


**Figure 1.** Visual bug summary, which we use to explain and illustrate the sea ice bug, but also propose as a template for easy bug communication in the future. a) Sea ice thickness in the high resolution simulations with the bug. b) Physical mechanism of the bug. c) Bug manifestation in the model code. Red highlights parts of the bug, which were fixed with the code highlighted in green. Here pseudo-code is shown to ease readability (see Appendix B for the actual differences). d) Sea ice thickness and ocean surface kinetic energy (KE, at ca. 200 $\mathrm{m}$ depth) in the high resolution simulation with a horizontal resolution of 5 km. The KE is a measure of the ocean surface velocity. It is calculated as $\mathrm{KE} = \frac{1}{2}\left(u^2 + v^2\right)$ where $u$ and $v$ is the ocean velocity in zonal and meridional direction, respectively. e) Same as d, but for the low resolution simulation with a horizontal resolution of 40 km.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


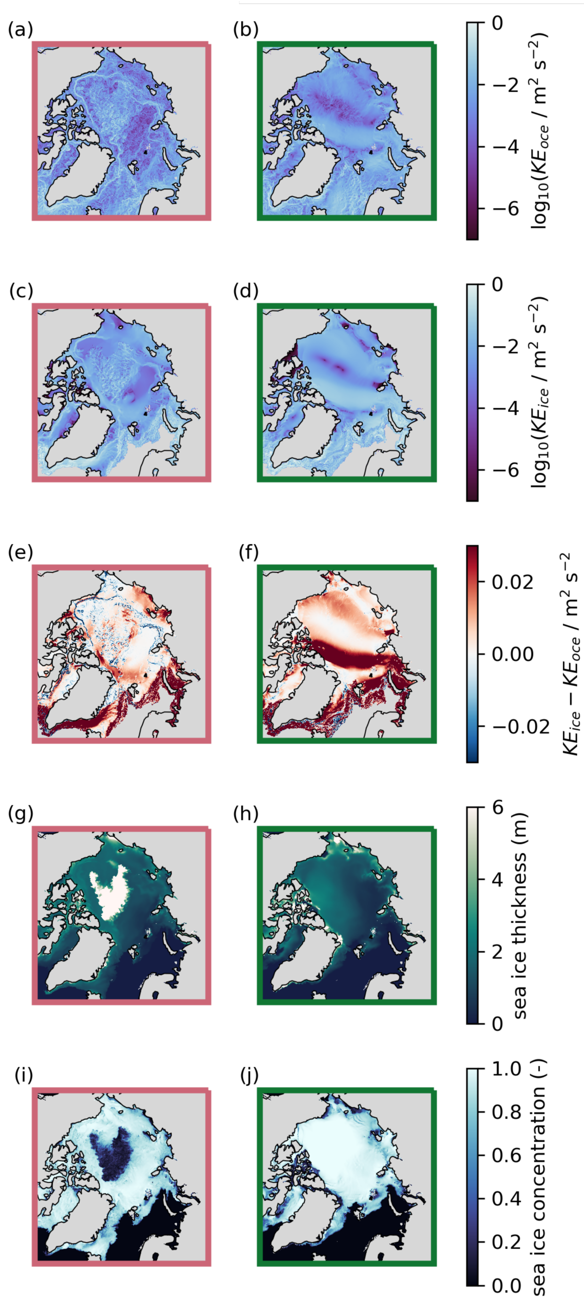


**Figure 2.** Simulation results from 20.04.2026 (as in Fig. 1), comparing the simulation with the bug (left column) and the one with the bugfix (right column). Displayed is the kinetic energy of the a) ocean surface, and b) sea ice; c) difference between the kinetic energies of sea ice and ocean; d) sea ice thickness; and e) sea ice concentration.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


### 3.1 The sea ice hole and its underlying physical mechanisms

The sea ice hole formation was characterized by a reduction in sea ice concentration and an increase in mean sea ice thickness at its boundaries. The regions of increased ice thickness were very narrow and propagated in a turbulent manner (see supplementary video). After the initialization of the model in 2020, the first formation of the sea ice hole started in August 2025. In the central Arctic and north of Greenland, multiple small regions of low ice concentration formed simultaneously. At their boundaries and also within the sea ice hole, very localized maxima in the sea ice thickness field were produced, because of the lead-close parameterization that adds newly formed ice to already existing ice within a grid cell. (see Fig. 2d) and e) for the juxtaposition of small concentrations and large sea ice thickness). These regions of low ice concentration grew with the turbulent movement of the sea ice. The sea ice hole grew until May 2026. During the melting phase in June, the regions of extensive ice thickness within the sea ice hole disappeared and the sea ice hole became a large region of open water covering half of the Arctic basin. By the end of August, the whole Arctic sea ice cover was gone. Starting in November 2026, the hole froze over again. In the years before and after the sea ice hole, the Arctic was always covered with sea ice in winter. The sea ice hole appeared only once during the 30 year simulation period. It is possible that other instances of a similar nature may exist in the simulation, but with a much more limited scope and impact. In other coupled configurations, which were similar in some respects but not identical to the present one, we observed a more frequent occurrence of the sea ice hole (see e.g. Fig. A1). It is noteworthy that similar, albeit smaller, holes were observed in uncoupled simulations, but only during the initialization phase that caused unrealistically strong ocean currents. Thus the bug was merely one of several potential mechanisms that can lead to strong ocean currents and trigger subsequent sea ice holes.

#### 3.1.1 Initial formation of the sea ice hole

The bug was found in the ocean dynamics where the sea ice to ocean momentum transfer is calculated. The total stress on the fluid ocean surface ($\tau_s$) is calculated from the relative contributions of the wind stress $\tau_w$ in regions of open water and from the ice-ocean stress $\tau_a$ in regions of sea ice cover:

$$\tau_s = \tau_a(1-A) + \tau_w A \tag{1}$$

where $A$ is the fraction of the grid cell covered with sea ice (and $1-A$ therefore the fraction of open water). The ice-ocean stress is calculated as:

$$\tau_w = \rho_0 C_d ||\mathbf{u_w} - \mathbf{v}||_2(\mathbf{u_w} - \mathbf{v}), \tag{2}$$

where $||\cdot||_2$ denotes the Euclidean norm, $\rho_0$ is a reference density of seawater, $C_d$ is the ice-ocean drag coefficient, $\mathbf{u_w}$ is the ocean velocity, and $\mathbf{v}$ is the sea ice velocity.

As described in Sec. 3.2, a flag in the source code unintentionally set the ice-ocean drag coefficient to zero in the ocean dynamics, while it was correctly applied in the sea ice dynamics. The total stress on the ocean surface was therefore reduced

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


to the wind stress

$$\tau_s = \tau_a(1 - A) \tag{3}$$

and underestimated in regions with high ice concentrations of $A \to 1$. In these regions, the ocean surface velocity should be slowed down by friction on the sea ice. Without this friction, the ocean did not lose momentum, resulting in excessive surface velocities, especially in eddies and boundary currents. Because momentum was correctly transferred to the sea ice, the ocean currents and eddies exerted too large stresses on the sea ice. These anomalous stress from the ocean, particularly from its eddy field, was the main factor that triggered the breakup of the sea ice.

#### 3.1.2 Effects of the spatial resolution on the sea ice hole formation

The sea ice hole only appeared in high-resolution simulations with a horizontal spatial resolution of 5 km. In the low-resolution simulation with a horizontal spatial resolution of 40 km, no sea ice holes formed. The simulation with 40 km resolution is too coarse for developing an eddy field. In contrast, we observed surprisingly many eddies in the simulation with 5 km resolution. A comparison of the ocean surface kinetic energy (KE) of a high-resolution simulation with a low-resolution simulation is shown in Fig. 1 d) and e). The increased eddy activity in the high-resolution simulation leads to substantial ocean-sea ice stresses on relatively small spatial scales. Fig. 2 shows that in the area of the hole, the kinetic energy of the ocean surface is comparable to that of the sea ice. With the bug fix, the kinetic energy is smoothed out in the ocean and in the sea ice. The enhanced stresses on small spatial scales in the 5 km simulation with the bug lead to strong sea ice deformations and are the primary cause of the sea ice hole formation.

#### 3.1.3 Lateral expansion of the sea ice hole

The sea ice hole was initially confined to a limited area, but it quickly expanded over nearly the entire Arctic basin. Several mechanisms probably contributed to the hole continuing to grow after it first appeared. The ice cover acts as an insulator between the ocean and the atmosphere, reducing the transfer of heat, momentum, and moisture. Once an open water area had formed, a positive feedback loop emerged, whereby enhanced heat fluxes out of the ocean result in increased near-surface air temperatures, which lead to further melting and thus larger ice-free areas. Over open water, the surface winds increased due to a reduction in surface roughness and also because of warmer sea surface temperatures, which decreased the vertical stability and enhanced downward mixing of momentum in the atmospheric boundary layer (Mioduszewski et al., 2018). The higher wind speeds caused a larger stress on the sea ice, leading to further sea ice deformation and subsequent formation of sea ice leads with open water. It is also possible that ocean surface eddy activity in regions of open water may be increased by higher wind speeds. Consequently, the total stress acting on the sea ice at the boundaries of the sea ice hole may be increased by these eddies. In addition, the warm air that accumulates over open water can be transported by the stronger winds in plumes over distances of several hundred kilometers from the sea ice hole (Gutjahr et al., 2016). These two effects represent a positive feedback loop that may act towards an enlargement of the sea ice hole.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


It is important to note that these feedbacks only occur in a coupled simulation where the atmospheric state is influenced by the ocean and the sea ice distribution. In addition to these atmospheric feedbacks, there are also oceanic feedbacks that lead to a reduction in sea ice cover. The albedo of sea water is lower than that of sea ice, which increased the absorption of solar radiation when the sea ice hole expanded during spring. The sea surface temperature increased locally, melting the nearby sea ice and reducing the formation of new sea ice. As new sea ice formed, the salinity and thus the density of the sea surface increased because most of the salt was not incorporated into the new sea ice. This increase in density at the water surface can lead to vertical mixing. Furthermore, in open water the wind stress can act directly on the ocean surface, thereby inducing vertical velocity shear and further mixing. The enhanced vertical mixing of temperature increased the volume of water that must be cooled to the freezing point before any new sea ice can form, thus preventing it from freezing over. Due to this delay, the sea ice cover is not restored immediately in open water areas, even though the air temperature is much lower than 0°C.

The feedback and processes explained in this section may lead to a further expansion and a maintenance of the sea ice hole. However, their relative contributions and whether there are other processes relevant to the development of the sea ice hole remain to be examined in detail.

### 3.2 Bug morphology

We have traced the evolution of the bug using the version control history recorded in the ICON code repository.

#### 3.2.1 Original implementation

In order to understand how the erroneous implementation came about, three switches are relevant (see Fig. 3):

- An option `stress_ice_zero` to disable sea ice induced stress on the ocean surface was introduced during the early stages of the sea ice model implementation, nearly ten years ago (in March 2015). The disabling was achieved by setting the contribution of the ice-ocean drag coefficient (`Cd_io`, see 3.1.1) to zero. The option was described by source code comments only. Presumably, it was added for testing. By default, the stress was (correctly) *not* set to zero (`stress_ice_zero = false`).
- `ice_dyn` controlled the computation of dynamical changes. In an early stage of development, coupling with the atmospheric model was incomplete, so the sea ice dynamics were disabled by default (`ice_dyn = 0`).
- `calc_ocean_stress` controls the stress calculation between ice and ocean.

As documented in a code comment, only one of the options `stress_ice_zero` and `calc_ocean_stress` was supposed to be set as they are contradictory. Nevertheless, at this point all three switches could be set independently without any consistency check.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024



**Figure 3.** Stresses in between atmosphere, ocean, and sea ice and how the three switches described in Sec. 3.2.1 act on them.

#### 3.2.2 Change of default behaviour and addition of erroneous consistency enforcement

In a follow-up code change, two consistency checks were added, with respect to the improved sea ice dynamics: If sea ice dynamics were enabled (`ice_dyn = 1`), the associated sea ice stress computation was re-enabled forcibly (`calc_ocean_stress = true`) and an appropriate warning was issued. As described above, in this case the sea ice induced stress should no longer be set to zero (`stress_ice_zero = false`). But while the second consistency check issues the correct warning, the subsequent code change erroneously enforced `stress_ice_zero = true` (see Fig. 1c). The code structure suggests that the second check was created by duplicating the first, and subsequently adapted incompletely. With this, the sea ice induced stress was always set to zero whenever the sea ice dynamics were enabled, ignoring any user supplied setting, thus counteracting the results from the sea ice stress computation. The erroneous change was not noticed for almost 10 years.

While the bug was mainly due to an oversight, we also identify advice for coding practice that could have made the construct easier to handle:

- Naming of the option: `stress_ice_zero` is ambiguous as it may refer to both ice-to-ocean and ocean-to-ice stress, and the meaning of zero is not entirely transparent. An explicit `disable_ice_stress_on_ocean` would have made the purpose much clearer.

- In general, intransparently altering an explicit user setting in favor of a more general option obliterates the code behavior for the model user, leading to unexpected results. If the two are mutually exclusive, the model should stop rather than issue a warning or second-guess the user's intent. Besides, user documentation for intent and usage of `stress_ice_zero` was missing.

- `Cd_io` is also used in the ocean to sea ice transfer of momentum and heat, but was not affected by `stress_ice_zero` in these contexts. From a physical point of view, a sea ice to ocean only modification of `Cd_io` seems rather idiosyncratic.

In addition, we question the practice of encapsulating testing or experimental procedures in if-statements to be accessed via the namelist. The technique is often used to avoid deleting code and to preserve backward compatibility, but it allows for erroneous configurations by design. In addition, as the code thereby becomes more complex, it increases the chance for oversights such as the one described above.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


### 3.2.3 Bugfix

To resolve the error, `stress_ice_zero` was again disabled by default. Also, the consistency enforcement was removed. It is still possible to both enable and disable `ice_dyn` and `stress_ice_zero` simultaneously, because that is currently still relied on in some of the model development tests. Both cases are flagged by specific warnings. Further assessment suggests that the use of `stress_ice_zero` is no longer relevant as the ICON sea ice model is quite mature by now. Once the development tests are updated accordingly, `stress_ice_zero` will be removed.

### 3.2.4 Appreciation of the error

The underlying error of documenting the logical switch but setting it incorrectly, is a classic mistake. According to McCauley et al. (2008), we classify it as a "blunder or botch", which they define as "a mental typo; you know what you wanted to write, but you wrote something else" (Knuth, 1989; McCauley et al., 2008). In the taxonomy developed by Avizienis (2004), our bug is a software flaw that occurs during development, is internal to the system, is human-made, affects the software, is non-malicious, non-deliberate, is introduced accidentally, and is permanent. The error is clearly not a model deficiency, but a mistake. While some implementation decisions were sub-optimal (see Sec. 3.2.2), the mistake itself was a slip in thinking. Thus, it could have been detected with careful code review. However, it cannot be attributed to a lack of programming skills, which is often lamented in the literature on scientific software development (Sundberg, 2009; Merali, 2010; Barnes and Jones, 2011). Once the bug was found, it was clear that it needed to be addressed. However, for some years the bug had remained undetected, because the simulations ran without problems, and at least in the low resolution setup, negative effects were too small to be noticed. On the contrary, as it served to open up leads in the winter, it increased formation of multi-year ice (see Sec. 3.3). As far as using the model as a tool for thinking and making projections is concerned, one could argue that in the low resolution setup the bug was insignificant. In the high resolution setup, it produced unphysical and large effects. It is this elusiveness of the bug that leads us to classify it as part of a new species: resolution-dependent bugs, whose effect depends on the resolution. They are present at all resolutions, but become important and detectable only in high resolution setups (the reverse is also conceivable). While the experience of finding bugs when moving to a new model configuration or extreme behaviour is not new, resolution-dependent bugs are a particularly relevant subclass that deserves emphasis and attention as climate models are increasingly applied in km scale setups.

## 3.3 How did the bug and bugfix affect the ocean and sea ice?

Fixing the sea ice bug had immediate positive effects on the Arctic ice cover. Even after running several tests with a total length of a couple of years, there was not a single occurrence of the sea ice hole anymore, in contrast to earlier experiments, where the ice hole occurred at an irregular frequency of less than five years. In particular, Fig. A1 shows the results of a simulation where the fix was applied in simulation year 1957. The sea ice hole appeared once more in 1958 and then never again until 1981. In addition, we observed a more pronounced occurrence of sea ice leads after the bug fix. Additionally, we found effects on the Arctic ocean circulation in all years, not just during years with the sea ice hole. While we observed a far too strong eddy

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


field under the ice before the bug fix, the situation changed after the sea ice bug was fixed. In particular, north of the Canadian archipelago, there were a large number of coherent eddies, which completely disappeared after the bug fix, and a coherent boundary current emerged.

Although the less eddying subsurface circulation, the increased frequency of sea ice leads, and the absence of the sea ice hole were all positive improvements, we also recognized that the Antarctic sea ice thickness became too thin. This feature is most likely an effect of reduced sea ice formation due to the absence of artificial holes as a result of the Antarctic Slope Current slowing down after the bug fix. However, our tuning choices for albedo, lead-close parameters, mixing, etc. were made for the false situation of a stronger upper-ocean circulation. Here, we are observing a common aspect of climate modeling, where some errors are masked by certain tuning choices. This reduced sea ice thickness is thus forcing us to reconsider some of our previous tuning choices of e.g. sea ice albedo, which we are currently working on.

Although the too strong eddy activity was most likely one important contributor for the sea ice hole to appear, there are several other mechanisms that likely contribute to the occurrence of the sea ice hole. One key aspect is the coupling of an atmosphere model which allows the above positive feedback mechanisms (see Sec. 3.1.3) to be active. We also observe a drift in the temperature and salinity fields (not shown) in the central Arctic. The warmer temperatures might also have contributed for the initialization of the hole. However, the sea ice hole only occurs infrequently and so far, we were not able to identify the initial cause of the hole.

Further open questions remain:

- Why did the sea ice holes occur with such an unpredictable frequency?
- Did vertical eddy heat fluxes have an influence on the formation of the sea ice hole?
- How did unrealistic currents affect the appearance of the sea ice hole?

Answers to those questions might be sought for heuristic reasons. However, as this manuscript focuses on bug communication and conceptualization, we deem a detailed analysis of the processes and feedbacks at play as out of scope.

Regarding previously published data and manuscripts, all model experiments using the ocean component of ICON are affected, as the discussed bug was present since the early development of ICON. Although the effect of the bug on coarse-resolution simulations is noticeable in retrospect, it is generally small. Furthermore, the effect is only directly relevant at high latitudes where sea ice is present, but might of course have indirect effects on other Earth system components via teleconnections. We reviewed the list of potentially affected papers that used the Sapphire configuration of the ICON model. However, most of these studies did not focus on sea ice or high latitudes and were therefore only indirectly affected by the bug, which we believe did not affect their main conclusions (Mauritsen et al., 2022; Hohenegger et al., 2023b; Bishnoi et al., 2024). Only two studies used a high-resolution simulation and focused on high latitudes. The papers discuss the effect of a katabatic storm and of polar lows on water mass transformation in the Arctic and subpolar North Atlantic (Gutjahr et al., 2022; Gutjahr and Mehlmann, 2024). While we have not investigated the bug's influence on sea ice conditions in these simulations, we consider the effect to be secondary to the extreme forcing from the storms and polar lows, which dominated the sea ice drift. In addition,

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


the studies focused on processes over short timescales of a few days to about three months. Furthermore, studies looking for mechanisms or process understanding are less threatened by bugs in general, as detailed in Sec. 5.

## 4 Bug finding experience

The following section presents a detailed recounting of the investigative process undertaken with regard to the sea ice bug during the subsequent nextGEMS hackathon. By highlighting *key features* of the process, we aim to contribute to the discourse on how debugging scientific software is done and can be improved. Aranda and Venolia (2009, Table 5) identified the **stakeholders' goals** during the lifespan of a bug, which we use to structure our account.

Since we are drawing conclusions from the peculiarities of the nextGEMS 4th hackathon (March 2024) setting, it is necessary to describe it briefly. The nextGEMS project develops and evaluates two km-scale resolution models, ICON and IFS/FESOM (Integrated Forecasting System/Finite Element Sea ice–Ocean Model, Rackow et al. (2024)). The project comprises 26 partners, with the Max Planck Institute for Meteorology leading the development of ICON Sapphire. Some meetings of the nextGEMS project take the form of hackathons, which means that there are few presentations and instead participants work with the model data on site. Importantly, this is an environment where collaboration is fostered and there is dedicated time to work on a specific project.

To tell the story of the bug, we identified the people involved and discussed what happened. The authors of this article represent the people most involved, thus we can get insight far beyond the electronic record, which is likely incomplete for bugs (Aranda and Venolia, 2009). We described in detail how the bug was noticed, investigated, found, understood and fixed. We have identified the circumstances and behaviors that are critical to history of the bug.

**Discovery** The Arctic sea ice holes, the most visible manifestation of the bug, had appeared repeatedly but sporadically in high resolution ICON simulations that ran for at least several years in coupled mode. These sea ice holes had previously been attempted to be addressed by tuning lead-close parameters, time-stepping for sea-ice rheology, etc. but to no avail and were thus particularly visible in the simulations prepared for the hackathon. During the opening presentations, the sea ice holes were shown repeatedly with *striking visualizations* (see Fig. 1a). Thus, there was *public pressure* on the sea ice and ocean team to address this peculiar model behaviour. The hackathon provided the fuel for a concerted effort to further investigate and resolve the issue of the sea ice holes, which we continue and communicate by writing this article. The investigation fits with the finding of Di Franco et al. (2017) that bugs seem to be found "almost by chance". Nobody was systematically testing for bugs of this kind, but the visualizations of model output alerted us. Clearly, social processes were involved in pushing the team to address the issue.

**Assignment of ownership** Working topics were announced in the opening session of the hackathon, where one of the authors (O. Gutjahr) declared that he would work on the bug during the week, and anyone interested was invited to join him and contribute.

The **search** for the appropriate knowledge, resources and skills was made much easier by the hackathon setup, which allowed *easy assembling of the team*, and many other scientists and scientific programmers could be asked for help. Bug fixing requires

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


domain knowledge (Di Franco et al., 2017), so in our case sea ice and ocean scientists were involved, as well as experienced scientific programmers, all with years of experience in working with the model.

**Diagnosis** The team collected possible causes of the sea ice holes and split up to investigate them. Specific hypotheses were, among others, that the rheology was not converging, or that the tuning of the model was faulty. Because tuning attempts previous to the hackathon failed to solve the problem, the team approached the problem from a physics perspective. Indications found during the investigation were that

- the sea ice holes only occurred where there were eddies.
- in plots, the ocean eddies and the ice drift fit perfectly together, while the physics would suggest that the ocean slows the ice, and that the wind drives the ice.
- the eddy field looked different when comparing ICON to FESOM. In this case it helped to have *multiple models* to compare against each other.

The clues we found led us to investigate the part of the source code where sea ice and ocean are coupled, which quickly led us to identify the problem (see Sec. 3). Importantly, producing the decisive *visualizations* and noticing the clues was the result of two days of thinking about the problem.

In terms of debugging tactics identified by Liu and Coblenz (2022, Table 2), the team employed reading the code, requesting help and change impact analysis, in the tuning attempts as well as for the final fix. Importantly, the switch to the debugging tactic of reading the code was a result of the clues listed above, confirming the debugging model put forward by Gould (1975). We classify the use of tuning before the investigation with physical understanding as the behaviour of "initially avoiding difficult sections of code" identified by Gould (1975), where in our case the complex code was avoided initially as a whole. In terms of coordination, the team was using *deep collaboration*, with occasional unexpected contributions from other scientists or programmers present at the hackathon (Aranda and Venolia, 2009, Table 4).

The hackathon setup in particular fostered finding the knowledge and interdisciplinary team, and the deep collaboration, because it provided the participants with a week of time scheduled to focus on the problem.

**Correction** After identifying the problem, the fixing of the bug was trivial. While the bug investigation was achieved within the week of the hackathon, testing of the bug fix and its effects continued afterwards.

**Awareness** The bug was communicated to participants at the end of the hackathon. This article aims to do the same for the entire ICON community. For judging whether a study is affected by the bug, we have described its effects and some criteria in Sec. 3.3.

## 5 Conclusions

We have conceptualized the sea ice bug documented in this paper as a resolution-dependent bug. By this we mean that while the bug was present in other, e.g. low-resolution configurations of the model, it was only detectable in the high-resolution setup. The reason was that the eddy velocities only became large enough at high resolution to have a noticeable effect on the

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


sea ice. Another example of a bug found during the ICON-Sapphire development that we ascribed to the resolution-dependent category, and that also appeared at the interface of two compartments, was a sign error in the coupling of ocean eddies and wind speed that caused them rotate into opposite directions.

Moving to km scales pushes models to their extremes, which can highlight erroneous behaviour. Of course, preparing the model or this new configuration also meant rewriting modules like the turbulence and mixing scheme, which introduced new errors. However, because the high-resolution model explicitly resolves more processes (for example, eddies are explicitly resolved rather than having their effects parameterized), it is also easier to subject it to physical reasoning, which helped in the debugging process.

Of course, the bug fix developed with the high-resolution model also benefits the low-resolution model. Martin et al. (2010) used this as a tactic, employing the NWP mode to identify the source of errors and develop fixes that then in turn benefited their climate model mode.

Could we have detected the bug earlier? The bug entered the model at a time when commits were coming in at a rapid rate. More stringent code and repository guidelines have since been introduced. In addition, there is a growing literature on tools for testing climate model code and finding bugs (Baker et al., 2015, 2016; Kim, 2016; Baker et al., 2017; Mahajan et al., 2019; Milroy et al., 2019; Ahn et al., 2021; Zeman and Schär, 2022), as well as individual examples of extensive documentation (Korkin, 2022). However, this bug was part of a legacy that is laborious to test in hindsight and for which it is difficult to provide guarantees. GCMs are complex pieces of code with many developers spread geographically and across time zones (Winsberg et al., 2014). Furthermore, we question whether this bug could have been found using testing procedures other than the scientific evaluation of the model results, which in the end pointed us to the error. It is notable that the effect on the low-resolution model was minimal if not favourable, suggesting that it was difficult to find the bug there. Our example serves to demonstrate a point about bugs in climate models made by Pipitone and Easterbrook (2012): code defects are either obvious (as in our high resolution case) or irrelevant (as arguably in the lower resolution case).

Recommendations we provide are in the form of observations of the circumstances that led us to undertake and successfully complete the debugging process (see Sec. 4): striking visualizations both in identifying and investigating the problem, public pressure, and the easy assembling of a team and deep collaborative work mode facilitated by the hackathon. During the hackathon, we engaged in a kind of sprint, where we had short daily meetings and a self organizing team. So one could argue that we were following some agile software development practices, but many of the formalized parts of that framework were missing from our approach (Sletholt et al., 2011). Importantly, the bug became obvious only irregularly every few years, which highlights the necessity for more than year long simulations also in km scale climate model setups.

It is clear that climate model code, as any human product, will continue to contain errors. Thus, in addition to improving control and testing mechanisms (Soergel, 2015), we must also find a constructive way of dealing with that fact scientifically and politically. This strategy could for example involve open communication of bugs as we suggest, and it needs to be developed at the institution and community level. As Stensrud (2007, pg. 394) put it, “numerical models are imperfect, so the key to success is how we deal with these imperfections.”

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


In general, the way scientists view bugs can be conceptualized in terms of the modeling vision they subscribe to. Combining concepts put forward as visions or epistemic communities by various authors (Shackley, 2001; Sundberg, 2009; Heymann and Hundebol, 2017; Heymann and Dahan Dalmedico, 2019), Proske (2023) summarizes three modeling visions: (1) the representative vision aims for the model to be a most exact representation of reality, with the more details the better; (2) in the heuristic vision, the model is a tool to investigate and generate understanding about the climate system; (3) in the predictive vision, the purpose of the model is to produce results that most closely resemble (future) observations. At least in the heuristic vision, the software implementation of a climate model can be viewed akin to 'evolving theories', and thus model imperfections can be accepted as inevitable (Easterbrook and Johns, 2009). In this view, a bug can be seen as an opportunity to generate knowledge, for example, by using the comparison between the fixed and the old model version simulations as a sensitivity experiment. As we highlighted in Sec. 3.3, a bug can even help uncover new scientific questions, linking to the popular notion of serendipity in science (Yaqub, 2018). This view highlights the interplay between the technical implementation and the physical science, and shows how "creating scientific software is often really the act of doing science itself" (Paine and Lee, 2017).

While a bug always poses a threat to the representative vision, the predictive vision may even first want to evaluate the effect of the bug fix on the results before deciding to implement it. Thus, as long as the model is adequate for its purpose (be it predictive or heuristic, Parker (2009)), the presence of unacknowledged or acknowledged errors does not disqualify the use of the model as a tool. In our case, the Arctic sea ice holes are clearly a threat to all three envisioned uses of the model, but they allow the use of the flawed simulations to generate understanding or to use the results from parts of the model that are less affected. Similarly, the bug does not render previous research results useless. As we have shown, differences in the results of the low resolution model are difficult to discern and in any case the bug had a direct influence only regionally.

Given the central role of climate models in climate science (Shackley et al., 1998; Edwards, 2001; Sundberg, 2007; Heymann et al., 2017; Heymann, 2019; Heymann and Dahan Dalmedico, 2019; Rödder et al., 2020) and their influence on policy and the public imagination (Mahony and Hulme, 2018; Heymann, 2010), how do bugs affect our confidence in climate models in general? A comprehensive philosophical treatment is beyond our expertise and scope, but we would like to leave the reader with some food for thought: climate scientists, and especially code developers, know that numerical models are "valuable, yet flawed, tools" (Stensrud, 2007, pg. 394). However, our confidence in their results is not only based on the false assumption that they are error-free (see e.g. Winsberg (2006); Knutti (2008b, a); Baumberger et al. (2017); Winsberg (2018); Palmer and Stevens (2019)). Heymann (2020) shows that, historically, their "epistemic justification was not rooted in the realism and mathematical rigour (...) but in the encouraging results". Mastroianni (2022) put forward the case that science is a strong-link problem, meaning that "progress depends on the quality of our best work". It is thus not the role of scientific review and model development to eradicate all mistakes, but rather to build on the strong ideas and results that survive in scientific discourse. Thus we can accept bugs as part of climate models and insist that we can still make scientific progress. We should be wary of promises to regard (high-resolution) models as perfect tools in favor of a more nuanced view: they are not perfect, but they offer exciting opportunities for investigation and understanding.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


### 5.1 Open bug communication

The case for change that we want to make is towards an open communication of bugs, their findings, and their effects, similar to the one Blanchard and Dũng (2024) made for publishing refutations. Our Fig. 1 aims to serve as a template that facilitates accessible and concise communication. Of course that template needs to be adapted to every case. In particular we realize that the bug presented in this case may present an easy example since it at its core involved only one wrongly set logical switch. Other bugs may be more involved, but that makes their communication and the summary of their core even more necessary.

More open communication can serve to increase trust, by the public but also within the scientific community. As Livingstone (2003, pg. 178) points out, "scientific knowing is an inescapably social phenomenon involving judgements about the integrity of people and their practices." In practice, open documentation is helpful not only for developers, but also for users of the climate model data. The latter need to know not only the intended function, but also the defects of a model in order to interpret its results. In addition, such publications can make explicit the art of debugging, which is tacit knowledge (Polanyi and Nye, 2015) that is expected of scientific programmers but not trained in scientific education. Also it relates to concepts of diversity and inclusion: Model data are provided openly in order to make them available to model users globally. However, if these users have no ties to the institution producing the model data, they consequently lack access to information on bugs contained in the model producing the data. As we have discussed above, published literature rarely treats bugs in more than a passing mention. Uncommunicated bugs become a threat for scientific quality: model data stays in use for far longer than any current model version where a bug may be fixed in the next version, so scientists interpreting model data need to have access to bug documentation. If models' data is open access, their faults need to be as well. On the side of model developers, as of now debugging model code is not funded or rewarded officially, and public documentation could at least serve to deliver recognition. In order for developers to make time for this work, bug documentation needs to become a norm with an established procedure (Kim and Zhang, 2015), similarly to how journals have enforced the open publication of datasets and software.

An open question is how to organise this bug communication. We can imagine either a database per model (for ICON, for example, that could be added to the existing website (https://www.icon-model.org/)), or a new paper type in a journal like GMD. The bug reports ideally would be integrated with the development software already in place, for example coming out of closed issues in GitLab. Doubtlessly, one would need to add attributes and metadata to make such a database serve its purpose of being accessible. Inclusion of bugs could be limited to those that have a detectable effect on model results. The elaboration and implementation of these ideas is up to the research community. With this paper, we make a start and hope to have provided a convincing example of the value and feasibility of open bug documentation.

*Code and data availability.* The ICON model is published in ICON partnership (DWD and MPI-M and DKRZ and KIT and C2SM) (2024) under a BSD 3-clause license. The low-resolution model version employed in this study corresponds to that release. We have published the high-resolution model versions (with the bugfix patch) as well as the plotting scripts and data used for the Figures in this study at Proske et al. (2024). The original report for the sea ice bug was filed at https://gitlab.dkrz.de/icon/icon-mpim/-/issues/86.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


*Video supplement.* Videos of the ocean kinetic energy and sea ice equivalent thickness are provided at Proske et al. (2024).

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


## Appendix A: EERIE Cycle 2 simulations

EERIE (European Eddy-Rich ESMs) is an EU-project whose objective is to understand the role of ocean eddies on the coupled climate system on longer-term timescales (multi-decadal to centennial). The model resolution employed (10 km atmosphere, 5 km ocean) is the same as the one used in this study. However, the length of the model integration for EERIE is much longer and hence the issue of sea ice holes was initially detected in EERIE simulations (50 years spin-up and 40 years control). In essence, the EERIE cycle 2 setup is the same as the nextGEMS cycle 4 setup used in this study (see Sec. 2.1), with the exception of being a control simulation using constant 1950-forcing. As shown in Fig. A1, for the first 7-years of integration while the bug fix was not implemented yet, the issue of sea ice holes occurred 3 times. The bug fix was implemented on year 8, and with the exception of year 9, the sea ice hole issue has not appeared again (see Fig. A1).

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


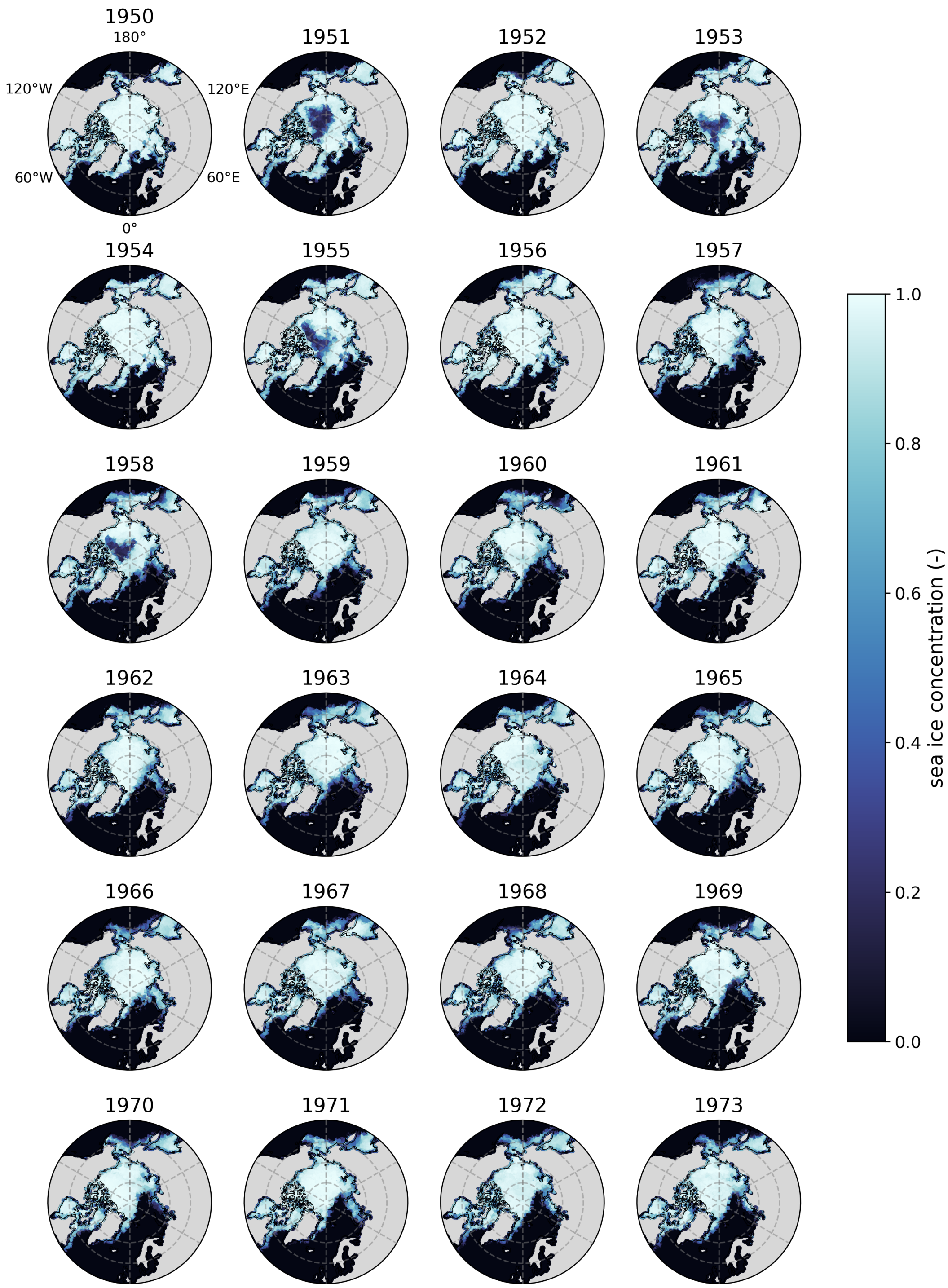


**Figure A1.** Northern Hemispheric sea ice concentration in March. The bugfix was implemented in the EERIE cycle 2 simulation from 1957 onward. Besides one single occurrence in March 1958, the sea ice holes have disappeared after implementation of the bug fix.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


## Appendix B: Code changes from the bug fix

The following shows the sea ice namelist code, with changes highlighted as differences between the code state before (red) and after (green) the bugfix ([. . . ] signifies code ommitted for the sake of brievity and clarity).

```
! namelist setup for the sea-ice model
!
!
! ICON
!
! ---------------------------------------------------------------
! Copyright (C) 2004-2024, DWD, MPI-M, DKRZ, KIT, ETH, MeteoSwiss
! Contact information: icon-model.org
!
! See AUTHORS.TXT for a list of authors
! See LICENSES/ for license information
! SPDX-License-Identifier: BSD-3-Clause
! ---------------------------------------------------------------

MODULE mo_sea_ice_nml

  USE mo_kind,                ONLY: wp

  [...]

  LOGICAL, PUBLIC :: use_constant_tfreez         = .TRUE.   !  constant freezing temperature for ocean water (=Tf)
  LOGICAL, PUBLIC :: use_IceInitialization_fromTemperature = .true.
-  LOGICAL, PUBLIC :: stress_ice_zero             = .TRUE.   !  set stress below sea ice to zero
+  LOGICAL, PUBLIC :: stress_ice_zero             = .FALSE.  !  set stress below sea ice to zero
  LOGICAL, PUBLIC :: use_calculated_ocean_stress = .FALSE.  !  calculate ocean stress instead of reading from OMIP
  LOGICAL, PUBLIC :: use_no_flux_gradients       = .TRUE.   !  simplified ice_fast without flux gradients


CONTAINS
  !>
  !!
  SUBROUTINE read_sea_ice_namelist(filename)

  [...]

    !-----------------------------------------------------------------
    ! Sanity Check
    !-----------------------------------------------------------------

    [...]

    IF (i_ice_dyn < 0 .OR. i_ice_dyn > 2) THEN
      CALL finish(TRIM(routine), 'i_ice_dyn must be either 0, 1 or 2.')
    END IF

-    IF (i_ice_dyn > 0 ) THEN
-      IF (i_ice_dyn == 1 ) THEN
```

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


```
-         CALL message(TRIM(routine), 'WARNING: i_ice_dyn is 1 - BUT SEA ICE DYNAMICS INCLUDE ERRORS')
-       ENDIF
-       ! TODO: This can be changed when we start advecting T1 and T2
-   !   CALL message(TRIM(routine), 'WARNING: i_ice_therm set to 1 because i_ice_dyn is 1')
-   !   i_ice_therm = 1   !  no Winton thermodynamics allowed, switched off by default
-
-       ! When using routine ice_ocean_stress, ocean stress below sea ice is considered accordingly
-       CALL message(TRIM(routine), 'WARNING: stress_ice_zero=FALSE because i_ice_dyn is 1 or 2')
-       stress_ice_zero = .TRUE.
+     IF (i_ice_dyn > 0 .AND. stress_ice_zero) THEN
+       CALL message(TRIM(routine), '---')
+       CALL message(TRIM(routine), 'WARNING: sea-ice dynamics are active but stress on ocean is disabled')
+       CALL message(TRIM(routine), 'WARNING: (sea_ice_nml.stress_ice_zero is true)')
+       CALL message(TRIM(routine), '---')
+     ENDIF
+     IF (i_ice_dyn == 0 .AND. .NOT. stress_ice_zero) THEN
+       CALL message(TRIM(routine), '---')
+       CALL message(TRIM(routine), 'WARNING: sea-ice dynamics are inactive but stress on ocean is enabled')
+       CALL message(TRIM(routine), 'WARNING: (sea_ice_nml.stress_ice_zero is false)')
+       CALL message(TRIM(routine), '---')
+     ENDIF

      [...]

  END SUBROUTINE read_sea_ice_namelist

END MODULE mo_sea_ice_nml
```

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


*Author contributions.* UP conceptualized the manuscript and administered the project. NB, JPG, HH, DP, KHW and OG investigated the bug. All authors wrote parts of the original draft and contributed to the review and editing of the manuscript.

*Competing interests.* The authors declare no competing interests.

*Acknowledgements.* The ideas in this manuscript were aided by discussions with the Atmospheric Chemistry and Aerosols Group at the Institute for Climate and Atmospheric Science at Leeds University, the nextGEMS and EERIE consortium, participants of the nextGEMS cycle 4 hackathon, and members of the Max Planck Institute for Meteorology in Hamburg, which we are grateful for. In particular we thank Johann Jungclaus, Rene Redler, Bjorn Stevens, Sabine Undorf, Veronika Gayler, Hauke Schmidt, Tobias Stacke and Florian Ziemen. UP was funded by the Swiss National Science foundation under grant number 217899. OG was funded by the European Union as part of the EPOC project (Explaining and Predicting the Ocean Conveyor) under call HORIZON-CL6-2021-CLIMATE01. This project has further received funding from the European Union's Horizon 2020 research program under Grant Agreement No 101003470 and 101081383 via the nextGEMS and EERIE projects respectively. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them. This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project IDs bm1235, bm1313, bm1344, k203123 and mh0033.

## References


Ahn, D. H., Baker, A. H., Bentley, M., Briggs, I., Gopalakrishnan, G., Hammerling, D. M., Laguna, I., Lee, G. L., Milroy, D. J., and Vertenstein, M.: Keeping Science on Keel When Software Moves, Communications of the ACM, https://doi.org/10.1145/ 3382037, 2021.

Aranda, J. and Venolia, G.: The Secret Life of Bugs: Going Past the Errors and Omissions in Software Repositories, in: Proceedings of the 31st International Conference on Software Engineering, https://doi.org/10.1109/ICSE.2009.5070530, 2009.

Avizienis, A.: Basic Concepts and Taxonomy of Dependable and Secure Computing, IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, 1, 2004.

Baker, A. H., Hammerling, D. M., Levy, M. N., Xu, H., Dennis, J. M., Eaton, B. E., Edwards, J., Hannay, C., Mickelson, S. A., Neale, R. B., Nychka, D., Shollenberger, J., Tribbia, J., Vertenstein, M., and Williamson, D.: A New Ensemble-Based Consistency Test for the Community Earth System Model (pyCECT v1.0), Geosci. Model Dev., 2015.

Baker, A. H., Hu, Y., Hammerling, D. M., Tseng, Y.-h., Xu, H., Huang, X., Bryan, F. O., and Yang, G.: Evaluating Statistical Consistency in the Ocean Model Component of the Community Earth System Model (pyCECT v2.0), Geosci. Model Dev., 2016.

Baker, A. H., Milroy, D. J., Hammerling, D. M., and Xu, H.: Quality Assurance and Error Identification for the Community Earth System Model, in: Proceedings of Correctness'17: First International Workshop on Software Correctness for HPC Applications (Correctness'17), https://doi.org/10.1145/3145344.3145491, 2017.

Barnes, N. and Jones, D.: Clear Climate Code: Rewriting Legacy Science Software for Clarity, IEEE Software, 28, 36–42, https://doi.org/10.1109/MS.2011.113, 2011.

Basili, V. R., Carver, J. C., Cruzes, D., Hochstein, L. M., Hollingsworth, J. K., Shull, F., and Zelkowitz, M. V.: Understanding the High-Performance-Computing Community: A Software Engineer's Perspective, IEEE Software, 25, 29–36, https://doi.org/10.1109/MS.2008.103, 2008.

Baumberger, C., Knutti, R., and Hirsch Hadorn, G.: Building Confidence in Climate Model Projections: An Analysis of Inferences from Fit, WIREs Climate Change, 8, https://doi.org/10.1002/wcc.454, 2017.

Bethke, I., Wang, Y., Counillon, F., Keenlyside, N., Kimmritz, M., Fransner, F., Samuelsen, A., Langehaug, H., Svendsen, L., Chiu, P.-G., Passos, L., Bentsen, M., Guo, C., Gupta, A., Tjiputra, J., Kirkevåg, A., Olivié, D., Seland, Ø., Vågane, J. S., Fan, Y., and Eldevik, T.: NorCPM1 and Its Contribution to CMIP6 DCPP, Geosci. Model Dev., 2021.

Bishnoi, A., Stein, O., Meyer, C. I., Redler, R., Eicker, N., Haak, H., Hoffmann, L., Klocke, D., Kornblueh, L., and Suarez, E.: Earth System Modeling on Modular Supercomputing Architecture: Coupled Atmosphere–Ocean Simulations with ICON 2.6.6-Rc, Geoscientific Model Development, 17, 261–273, https://doi.org/10.5194/gmd-17-261-2024, 2024.

Blanchard, E. and Dũng, L. T.: Papers Have Bugs — What Is to Be Done?, Undone Computer Science Conference, 2024.

Calotoiu, A., Hoefler, T., Poke, M., and Wolf, F.: Using Automated Performance Modeling to Find Scalability Bugs in Complex Codes, in: SC '13: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, https://doi.org/10.1145/2503210.2503277, 2013.

Carver, J. C., Kendall, R. P., Squires, S. E., and Post, D. E.: Software Development Environments for Scientific and Engineering Software: A Series of Case Studies, th International Conference on Software Engineering, 2007.

Chowdhury, I. and Zulkernine, M.: Can Complexity, Coupling, and Cohesion Metrics Be Used as Early Indicators of Vulnerabilities?, in: SAC '10: Proceedings of the 2010 ACM Symposium on Applied Computing, https://doi.org/10.1145/1774088.1774504, 2010.

Danilov, S., Wang, Q., Timmermann, R., Iakovlev, N., Sidorenko, D., Kimmritz, M., Jung, T., and Schröter, J.: Finite-Element Sea Ice Model (FESIM), version 2, Geoscientific Model Development, 8, 1747–1761, https://doi.org/10.5194/gmd-8-1747-2015, 2015.

Depaz, P.: The Role of Aesthetics in Understanding Source Code, Ph.D. thesis, Université Sorbonne Nouvelle, 2023.

Di Franco, A., Guo, H., and Rubio-Gonzalez, C.: A Comprehensive Study of Real-World Numerical Bug Characteristics, in: 2017 32nd IEEE/ACM International Conference on Automated Software Engineering (ASE), pp. 509–519, IEEE, Urbana, IL, ISBN 978-1-5386-2684-9, https://doi.org/10.1109/ASE.2017.8115662, 2017.

Easterbrook, S. M. and Johns, T. C.: Engineering the Software for Understanding Climate Change, Computing in science & engineering, 2009.

Edwards, P. N.: Representing the Global Atmosphere, in: Changing the Atmosphere, The MIT Press, ISBN 978-0-262-27981-9, https://doi.org/10.7551/mitpress/1789.003.0005, 2001.

Gould, J. D.: Some Psychological Evidence on How People Debug Computer Programs, International Journal of Man-Machine Studies, 7, https://doi.org/10.1016/S0020-7373(75)80005-8, 1975.

Graff, L. S., Iversen, T., Bethke, I., Debernard, J. B., Seland, Ø., Bentsen, M., Kirkevåg, A., Li, C., and Olivié, D. J. L.: Arctic Amplification under Global Warming of 1.5 and 2 °C in NorESM1-Happi, Earth System Dynamics, https://doi.org/10.5194/esd-10-569-2019, 2019.

Gutjahr, O. and Mehlmann, C.: Polar Lows and Their Effects on Sea Ice and the Upper Ocean in the Iceland, Greenland, and Labrador Seas, Journal of Geophysical Research: Oceans, 129, e2023JC020 258, https://doi.org/https://doi.org/10.1029/2023JC020258, e2023JC020258 2023JC020258, 2024.

Gutjahr, O., Heinemann, G., Preußer, A., Willmes, S., and Drüe, C.: Quantification of ice production in Laptev Sea polynyas and its sensitivity to thin-ice parameterizations in a regional climate model, The Cryosphere, 10, 2999–3019, https://doi.org/10.5194/tc-10-2999-2016, 2016.

Gutjahr, O., Jungclaus, J. H., Brüggemann, N., Haak, H., and Marotzke, J.: Air-Sea Interactions and Water Mass Transformation During a Katabatic Storm in the Irminger Sea, Journal of Geophysical Research: Oceans, 127, e2021JC018 075, https://doi.org/10.1029/2021JC018075, 2022.

Hajek, E. A. and Wolinsky, M. A.: Simplified Process Modeling of River Avulsion and Alluvial Architecture: Connecting Models and Field Data, Sedimentary Geology, 257-260, 1–30, https://doi.org/10.1016/j.sedgeo.2011.09.005, 2012.

Hatton, L.: The T-Experiments: Errors in Scientific Software, Quality of Numerical Software, 1997.

Hatton, L.: The Chimera of Software Quality, Computer, 40, 104–103, https://doi.org/10.1109/MC.2007.292, 2007.

Heymann, M.: The Evolution of Climate Ideas and Knowledge, WIREs Climate Change, 1, 581–597, https://doi.org/10.1002/wcc.61, 2010.

Heymann, M.: The Climate Change Dilemma: Big Science, the Globalizing of Climate and the Loss of the Human Scale, Regional Environmental Change, 19, 1549–1560, https://doi.org/10.1007/s10113-018-1373-z, 2019.

Heymann, M.: Knowledge Production with Climate Models: On the Power of a "Weak" Type of Knowledge, in: Weak Knowledge: Forms, Functions, and Dynamics, pp. 321–349, Campus Verlag, Frankfurt, ISBN 978-3-593-50977-8, 2020.

Heymann, M. and Dahan Dalmedico, A.: Epistemology and Politics in Earth System Modeling: Historical Perspectives, Journal of Advances in Modeling Earth Systems, 11, 1139–1152, https://doi.org/10.1029/2018MS001526, 2019.

Heymann, M. and Hundebol, N. R.: From Heuristic to Predictive: Making Climate Models into Political Instruments, in: Cultures of Prediction in Atmospheric and Climate Science, pp. 100–119, Routledge, ISBN 978-0-262-63219-5, 2017.

Heymann, M., Gramelsberger, G., and Mahony, M.: Key Characteristics of Cultures of Prediction, in: Cultures of Prediction in Atmospheric and Climate Science, pp. 18–41, Routledge, Taylor & Francis Group, ISBN 978-0-262-63219-5, 2017.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


Hibler, W. D.: A Dynamic Thermodynamic Sea Ice Model, Journal of Physical Oceanography, 9, 815–846, https://doi.org/https://doi.org/10.1175/1520-0485(1979)009<0815:ADTSIM>2.0.CO;2, 1979.

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., Bao, J., Bastin, S., Behravesh, M., Bergemann, M., Biercamp, J., Bockelmann, H., Brokopf, R., Brüggemann, N., Casaroli, L., Chegini, F., Datseris, G., Esch, M., George, G., Giorgetta, M., Gutjahr, O., Haak, H., Hanke, M., Ilyina, T., Jahns, T., Jungclaus, J., Kern, M., Klocke, D., Kluft, L., Kölling, T., Kornblueh, L., Kosukhin, S., Kroll, C., Lee, J., Mauritsen, T., Mehlmann, C., Mieslinger, T., Naumann, A. K., Paccini, L., Peinado, A., Praturi, D. S., Putrasahan, D., Rast, S., Riddick, T., Roeber, N., Schmidt, H., Schulzweida, U., Schütte, F., Segura, H., Shevchenko, R., Singh, V., Specht, M., Stephan, C. C., von Storch, J.-S., Vogel, R., Wengel, C., Winkler, M., Ziemen, F., Marotzke, J., and Stevens, B.: ICON-Sapphire: simulating the components of the Earth system and their interactions at kilometer and subkilometer scales, Geoscientific Model Development, 16, 779–811, https://doi.org/10.5194/gmd-16-779-2023, 2023a.

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., Bao, J., Bastin, S., Behravesh, M., Bergemann, M., Biercamp, J., Bockelmann, H., Brokopf, R., Brüggemann, N., Casaroli, L., Chegini, F., Datseris, G., Esch, M., George, G., Giorgetta, M., Gutjahr, O., Haak, H., Hanke, M., Ilyina, T., Jahns, T., Jungclaus, J., Kern, M., Klocke, D., Kluft, L., Kölling, T., Kornblueh, L., Kosukhin, S., Kroll, C., Lee, J., Mauritsen, T., Mehlmann, C., Mieslinger, T., Naumann, A. K., Paccini, L., Peinado, A., Praturi, D. S., Putrasahan, D., Rast, S., Riddick, T., Roeber, N., Schmidt, H., Schulzweida, U., Schütte, F., Segura, H., Shevchenko, R., Singh, V., Specht, M., Stephan, C. C., von Storch, J.-S., Vogel, R., Wengel, C., Winkler, M., Ziemen, F., Marotzke, J., and Stevens, B.: ICON-Sapphire: Simulating the Components of the Earth System and Their Interactions at Kilometer and Subkilometer Scales, Geoscientific Model Development, 16, 779–811, https://doi.org/10.5194/gmd-16-779-2023, 2023b.

Hook, D. and Kelly, D.: Testing for Trustworthiness in Scientific Software, in: 2009 ICSE Workshop on Software Engineering for Computational Science and Engineering, pp. 59–64, https://doi.org/10.1109/SECSE.2009.5069163, 2009.

Hook, D. A.: Using Code Mutation to Study Code Faults in Scientific Software, Ph.D. thesis, Queen's University, Kingston, Ontario, Canada, 2009.

Hunke, E. C. and Dukowicz, J. K.: An Elastic–Viscous–Plastic Model for Sea Ice Dynamics, Journal of Physical Oceanography, 27, 1849–1867, https://doi.org/https://doi.org/10.1175/1520-0485(1997)027<1849:AEVPMF>2.0.CO;2, 1997.

ICON partnership (DWD and MPI-M and DKRZ and KIT and C2SM): ICON release 2024.01, https://doi.org/10.35089/WDCC/IconRelease01, 2024.

Jöckel, P., Tost, H., Pozzer, A., Kunze, M., Kirner, O., Brenninkmeijer, C. A. M., Brinkop, S., Cai, D. S., Dyroff, C., Eckstein, J., Frank, F., Garny, H., Gottschaldt, K.-D., Graf, P., Grewe, V., Kerkweg, A., Kern, B., Matthes, S., Mertens, M., Meul, S., Neumaier, M., Nützel, M., Oberländer-Hayn, S., Ruhnke, R., Runde, T., Sander, R., Scharffe, D., and Zahn, A.: Earth System Chemistry Integrated Modelling (ESCiMo) with the Modular Earth Submodel System (MESSy) Version 2.5, Geosci. Model Dev., 2, 2016.

Kahre, M., Haberle, R., Wilson, R., Urata, R., Steakley, K., Brecht, A., Bertrand, T., Kling, A., Batterson, C., Hartwick, V., Harman, C., and Gkouvelis, L.: The NASA Ames Legacy Mars Global Climate Model: Radiation Code Error Correction and New Baseline Water Cycle Simulation, Icarus, 400, 115 561, https://doi.org/10.1016/j.icarus.2023.115561, 2023.

Kelly, D.: Scientific Software Development Viewed as Knowledge Acquisition: Towards Understanding the Development of Risk-Averse Scientific Software, The Journal of Systems and Software, 109, https://doi.org/10.1016/j.jss.2015.07.027, 2015.

Kelly, D. and Sanders, R.: Assessing the Quality of Scientific Software, First International Workshop on Software Engineering for Computational Science and Engineering, 2008.

https://doi.org/10.5194/egusphere-2024-3493
Preprint. Discussion started: 28 November 2024


Kendall, R., Fisher, D., Henderson, D., Carver, J. C., Mark, A., Post, D., Jr, C. E. R., Squires, S., and Microsystems, S.: Development of a Weather Forecasting Code:, IEEE, 2008.

Kim, Y.: KGEN: A Python Tool for Automated Fortran Kernel Generation and Verification, Procedia Computer Science, 80, https://doi.org/10.1016/j.procs.2016.05.466, 2016.

Kim, Y. and Zhang, P.: Understanding data sharing behaviors of STEM researchers: The roles of attitudes, norms, and data repositories, Library & Information Science Research, 37, 189–200, https://doi.org/10.1016/j.lisr.2015.04.006, 2015.

Kimmritz, M., Losch, M., and Danilov, S.: A comparison of viscous-plastic sea ice solvers with and without replacement pressure, Ocean Modelling, 115, 59–69, https://doi.org/https://doi.org/10.1016/j.ocemod.2017.05.006, 2017.

Knuth, D. E.: The Errors of TEX, Software - Practice and Experience, 19, 607–685, https://doi.org/10.1002/spe.4380190702, 1989.

Knutti, R.: Hotter or Not? Should We Believe Model Predictions of Future Climate Change?, Significance, 5, 159–162, https://doi.org/10.1111/j.1740-9713.2008.00320.x, 2008a.

Knutti, R.: Should We Believe Model Predictions of Future Climate Change?, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 366, 4647–4664, https://doi.org/10.1098/rsta.2008.0169, 2008b.

Korkin, S.: A Practical Guide to Writing a Radiative Transfer Code, Computer Physics Communications, 2022.

Korn, P.: Formulation of an unstructured grid model for global ocean dynamics, J. Comp. Physiol., 339, 525–552, https://doi.org/10.1016/j.jcp.2017.03.009, 2017.

Korn, P., Brüggemann, N., Jungclaus, J. H., Lorenz, S. J., Gutjahr, O., Haak, H., Linardakis, L., Mehlmann, C., Mikolajewicz, U., Notz, D., Putrasahan, D. A., Singh, V., von Storch, J.-S., Zhu, X., and Marotzke, J.: ICON-O: The Ocean Component of the ICON Earth System Model—Global Simulation Characteristics and Local Telescoping Capability, Journal of Advances in Modeling Earth Systems, 14, e2021MS002 952, https://doi.org/https://doi.org/10.1029/2021MS002952, 2022.

Kreyscher, M., Harder, M., Lemke, P., and Flato, G. M.: Results of the Sea Ice Model Intercomparison Project: Evaluation of sea ice rheology schemes for use in climate simulations, Journal of Geophysical Research: Oceans, 105, 11 299–11 320, https://doi.org/https://doi.org/10.1029/1999JC000016, 2000.

Larsen, L. G., Eppinga, M. B., Passalacqua, P., Getz, W. M., Rose, K. A., and Liang, M.: Appropriate Complexity Landscape Modeling, Earth-Science Reviews, 160, 111–130, https://doi.org/10.1016/j.earscirev.2016.06.016, 2016.

Liu, A. and Coblenz, M.: Debugging Techniques in Professional Programming, in: Plateau Workshop, https://doi.org/10.1184/R1/22277365.v1, 2022.

Livingstone, D. N.: Putting Science in Its Place: Geographies of Scientific Knowledge, Science.Culture, University of Chicago Press, Chicago, ISBN 978-0-226-48722-9, 2003.

Luther, F. M., Ellingson, R. G., Fouquart, Y., Fels, S., Scott, N. A., and Wiscombe, W. J.: Intercomparison of Radiation Codes in Climate Models (ICRCCM): Longwave Clear-Sky Results—A Workshop Summary, Bulletin American Meteorological Society, 69, 1988.

Mahajan, S., Evans, K. J., Kennedy, J. H., Xu, M., and Norman, M. R.: A Multivariate Approach to Ensure Statistical Reproducibility of Climate Model Simulations, in: PASC ’19, https://doi.org/10.1145/3324989.3325724, 2019.

Mahony, M. and Hulme, M.: Epistemic Geographies of Climate Change: Science, Space and Politics, Progress in Human Geography, 42, 395–424, https://doi.org/10.1177/0309132516681485, 2018.

Martin, G. M., Milton, S. F., Senior, C. A., Brooks, M. E., Ineson, S., Reichler, T., and Kim, J.: Analysis and Reduction of Systematic Errors through a Seamless Approach to Modeling Weather and Climate, Journal of Climate, 23, 5933–5957, https://doi.org/10.1175/2010JCLI3541.1, 2010.

Mastroianni, A.: The Rise and Fall of Peer Review, 2022.

Mauritsen, T., Redler, R., Esch, M., Stevens, B., Hohenegger, C., Klocke, D., Brokopf, R., Haak, H., Linardakis, L., Röber, N., and Schnur, R.: Early Development and Tuning of a Global Coupled Cloud Resolving Model, and Its Fast Response to Increasing CO2, Tellus A: Dynamic Meteorology and Oceanography, 74, 346–363, https://doi.org/10.16993/tellusa.54, 2022.

McCauley, R., Fitzgerald, S., Lewandowski, G., Murphy, L., Simon, B., Thomas, L., and Zander, C.: Debugging: A Review of the Literature from an Educational Perspective, Computer Science Education, 18, 67–92, https://doi.org/10.1080/08993400802114581, 2008.

McGraw, Z., Storelvmo, T., Polvani, L. M., Hofer, S., Shaw, J. K., and Gettelman, A.: On the Links Between Ice Nucleation, Cloud Phase, and Climate Sensitivity in CESM2, Geophysical Research Letters, 50, e2023GL105 053, https://doi.org/10.1029/2023GL105053, 2023.

Menard, C. B., Essery, R., Krinner, G., Arduini, G., Bartlett, P., Boone, A., Brutel-Vuilmet, C., Burke, E., Cuntz, M., Dai, Y., Decharme, B., Dutra, E., Fang, X., Fierz, C., Gusev, Y., Hagemann, S., Haverd, V., Kim, H., Lafaysse, M., Marke, T., Nasonova, O., Nitta, T., Niwano, M., Pomeroy, J., Schädler, G., Semenov, V. A., Smirnova, T., Strasser, U., Swenson, S., Turkov, D., Wever, N., and Yuan, H.: Scientific and Human Errors in a Snow Model Intercomparison, Bulletin of the American Meteorological Society, 102, E61–E79, https://doi.org/10.1175/BAMS-D-19-0329.1, 2021.

Mendez, M., Tinetti, F. G., and Overbey, J. L.: Climate Models: Challenges for Fortran Development Tools, in: 2014 Second International Workshop on Software Engineering for High Performance Computing in Computational Science and Engineering, pp. 6–12, IEEE, New Orleans, LA, USA, ISBN 978-1-4799-7035-3, https://doi.org/10.1109/SE-HPCCSE.2014.7, 2014.

Merali, Z.: Computational Science: ...Error, Nature, 467, 775–777, https://doi.org/10.1038/467775a, 2010.

Miller, G.: A Scientist's Nightmare: Software Problem Leads to Five Retractions, Science, 314, 2006.

Milroy, D. J., Baker, A. H., Hammerling, D. M., Kim, Y., Jessup, E. R., and Hauser, T.: Making Root Cause Analysis Feasible for Large Code Bases: A Solution Approach for a Climate Model, in: Proceedings of the 28th International Symposium on High-Performance Parallel and Distributed Computing, HPDC '19, pp. 73–84, Association for Computing Machinery, New York, NY, USA, ISBN 978-1-4503-6670-0, https://doi.org/10.1145/3307681.3325399, 2019.

Mioduszewski, J., Vavrus, S., and Wang, M.: Diminishing Arctic sea ice promotes stronger surface winds, Journal of Climate, 31, 8101–8119, 2018.

Müller, W., Früh, B., Korn, P., Potthast, R., Baehr, J., Bettems, J.-M., Bölöni, G., Brienen, S., Fröhlich, K., Helmert, J., Jungclaus, J., Köhler, M., Lorenz, S., Schneidereit, A., Schnur, R., Schulz, J.-P., Schlemmer, L., Sgoff, C., Pham, T., Pohlmann, H., Vogel, B., Vogel, H., Wirth, R., Zaehle, S., Zängl, G., Stevens, B., and Marotzke, J.: ICON: Towards vertically integrated model configurations for numerical weather prediction, climate prediction and projections, BAMs, submitted, 2024.

Newman, J.: Epistemic Opacity, Confirmation Holism and Technical Debt: Computer Simulation in the Light of Empirical Software Engineering, in: History and Philosophy of Computing: Third International Conference, HaPoC 2015 Pisa, Italy, October 8-11,2015 ; Revised Selected Papers, edited by Gadducci, F. and Tavosanis, M., no. 487 in IFIP Advances in Information and Communication Technology, Springer International Publishing AG, Cham, ISBN 978-3-319-47286-7 978-3-319-47285-0, 2016.

nextGEMS: nextGEMS | Next Generation Earth Modelling Systems, https://nextgems-h2020.eu/, 2024.

Paine, D. and Lee, C. P.: "Who Has Plots?": Contextualizing Scientific Software, Practice, and Visualizations, in: Proc. ACM Hum.-Comput. Interact., https://doi.org/10.1145/3134720, 2017.

Palmer, T. and Stevens, B.: The Scientific Challenge of Understanding and Estimating Climate Change, Proceedings of the National Academy of Sciences, 116, 24 390–24 395, https://doi.org/10.1073/pnas.1906691116, 2019.

Parker, W. S.: Confirmation and Adequacy-for-Purpose in Climate Modelling, Aristotelian Society Supplementary Volume, 83, 233–249, https://doi.org/10.1111/j.1467-8349.2009.00180.x, 2009.

Pipitone, J. and Easterbrook, S.: Assessing Climate Model Software Quality: A Defect Density Analysis of Three Models, Geosci. Model Dev., 2012.

Polanyi, M. and Nye, M. J.: Personal Knowledge: Towards a Post-Critical Philosophy, University of Chicago Press, Chicago, enlarged edition edn., ISBN 978-0-226-23262-1, 2015.

Proske, U.: Simplifying the Cloud Microphysics and Aerosol Representation of a Global Aerosol Climate Model, Doctoral Thesis, ETH Zurich, https://doi.org/10.3929/ethz-b-000646076, 2023.

Proske, U., Brüggemann, N., Gärtner, J. P., Gutjahr, O., Haak, H., Putrasahan, D., and Wieners, K.-H.: Data and scripts for the publication "A case for open communication of bugs in climate models", https://doi.org/10.5281/zenodo.14220612, 2024.

Rackow, T., Pedruzo-Bagazgoitia, X., Becker, T., Milinski, S., Sandu, I., Aguridan, R., Bechtold, P., Beyer, S., Bidlot, J., Boussetta, S., Dueben, P., Dutra, E., Forbes, R., Goessling, H. F., Hadade, I., Keeley, S., Kluft, L., Koldunov, N., Koldunov, A., Kölling, T., Mogensen, K., Quintino, T., Polichtchouk, I., Sármány, D., Streffing, J., Sützl, B., Takasuka, D., Tietsche, S., Valentini, M., Wedi, N., Zampieri, L., and Ziemen, F.: Multi-Year Simulations at Kilometre Scale with the Integrated 2 Forecasting System Coupled to FESOM2.5/NEMOv3.4, 2024.

Rahman, A., Bose, D. B., Barsha, F. L., and Pandita, R.: Defect Categorization in Compilers: A Multi-Vocal Literature Review, ACM Computing Surveys, 56, 2023.

Reick, C. H., Gayler, V., Goll, D., Hagemann, S., Heidkamp, M., Nabel, J. E. M. S., Raddatz, T., Roeckner, E., Schnur, R., and Wilkenskjeld, S.: JSBACH 3 - The land component of the MPI Earth System Model: documentation of version 3.2, Berichte zur Erdsystemforschung, 240, https://doi.org/10.17617/2.3279802, 2021.

Rödder, S., Heymann, M., and Stevens, B.: Historical, Philosophical, and Sociological Perspectives on Earth System Modeling, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2020MS002139, 2020.

Segal, J.: Models of Scientific Software Development, 2004.

Segal, J.: Software Development Cultures and Cooperation Problems: A Field Study of the Early Stages of Development of Software for a Scientific Community, Computer Supported Cooperative Work, https://doi.org/10.1007/s10606-009-9096-9, 2009.

Semtner, A. J.: A Model for the Thermodynamic Growth of Sea Ice in Numerical Investigations of Climate, Journal of Physical Oceanography, 6, 379–389, https://doi.org/https://doi.org/10.1175/1520-0485(1976)006<0379:AMFTTG>2.0.CO;2, 1976.

Shackley, S.: Epistemic Lifestyles in Climate Change Modelling, in: Changing the Atmosphere, MIT Press, ISBN 978-0-262-63219-5, 2001.

Shackley, S., Young, P., Parkinson, S., and Wynne, B.: Uncertainty, Complexity and Concepts of Good Science in Climate Change Modelling: Are GCMs the Best Tools?, Climatic Change, 38, 159–205, https://doi.org/10.1023/A:1005310109968, 1998.

Shaw, J., McGraw, Z., Bruno, O., Storelvmo, T., and Hofer, S.: Using Satellite Observations to Evaluate Model Microphysical Representation of Arctic Mixed-Phase Clouds, Geophysical Research Letters, https://doi.org/10.1029/2021GL096191, 2022.

Sletholt, M. T., Hannay, J., Pfahl, D., Benestad, H. C., and Langtangen, H. P.: A Literature Review of Agile Practices and Their Effects in Scientific Software Development, in: Proceedings of the 4th International Workshop on Software Engineering for Computational Science and Engineering, SECSE '11, pp. 1–9, Association for Computing Machinery, New York, NY, USA, ISBN 978-1-4503-0598-3, https://doi.org/10.1145/1985782.1985784, 2011.

Soergel, D. A. W.: Rampant Software Errors May Undermine Scientific Results, F1000 Research, https://doi.org/10.12688/f1000research.5930.2, 2015.

Stensrud, D. J.: Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models, Cambridge University Press, first edn., ISBN 978-0-521-86540-1 978-0-521-12676-2 978-0-511-81259-0, https://doi.org/10.1017/CBO9780511812590, 2007.

Storer, T.: Bridging the Chasm, ACM Computing Surveys, 50, 2017.

Sundberg, M.: Parameterizations as Boundary Objects on the Climate Arena, Social Studies of Science, 37, 473–488, https://doi.org/10.1177/0306312706075330, 2007.

Sundberg, M.: The Everyday World of Simulation Modeling: The Development of Parameterizations in Meteorology, Science, Technology, & Human Values, 34, 162–181, https://doi.org/10.1177/0162243907310215, 2009.

Tallapragada, V., Kieu, C., Kwon, Y., Trahan, S., Liu, Q., Zhang, Z., and Kwon, I.-H.: Evaluation of Storm Structure from the Operational HWRF during 2012 Implementation, MONTHLY WEATHER REVIEW, 142, 2014.

Tucker, G. E., Hutton, E. W. H., Piper, M. D., Campforts, B., Gan, T., Barnhart, K. R., Kettner, A. J., Overeem, I., Peckham, S. D., McCready, L., and Syvitski, J.: CSDMS: A Community Platform for Numerical Modeling of Earth Surface Processes, Geosci. Model Dev., 2022.

Wilson, G. V.: Where's the Real Bottleneck in Scientific Computing?, American Scientist, 94, 2006.

Winsberg, E.: Models of Success Versus the Success of Models: Reliability without Truth, Synthese, 152, 1–19, https://doi.org/10.1007/s11229-004-5404-6, 2006.

Winsberg, E., Huebner, B., and Kukla, R.: Accountability and Values in Radically Collaborative Research, Studies in History and Philosophy of Science Part A, 46, 16–23, https://doi.org/10.1016/j.shpsa.2013.11.007, 2014.

Winsberg, E. B.: Philosophy and Climate Science, Cambridge University Press, Cambridge, United Kingdom ; New York, NY, ISBN 978-1-107-19569-1 978-1-316-64692-2, 2018.

Yaqub, O.: Serendipity: Towards a Taxonomy and a Theory, Research Policy, 47, 169–179, https://doi.org/10.1016/j.respol.2017.10.007, 2018.

Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-Hydrostatic) Modelling Framework of DWD and MPI-M: Description of the Non-Hydrostatic Dynamical Core, Quarterly Journal of the Royal Meteorological Society, 141, 563–579, https://doi.org/10.1002/qj.2378, 2015.

Zeman, C. and Schär, C.: An Ensemble-Based Statistical Methodology to Detect Differences in Weather and Climate Model Executables, Geoscientific Model Development, 15, 3183–3203, https://doi.org/10.5194/gmd-15-3183-2022, 2022.

Zhu, J., Otto-Bliesner, B. L., Brady, E. C., Gettelman, A., Bacmeister, J. T., Neale, R. B., Poulsen, C. J., Shaw, J. K., McGraw, Z. S., and Kay, J. E.: LGM Paleoclimate Constraints Inform Cloud Parameterizations and Equilibrium Climate Sensitivity in CESM2, Journal of Advances in Modeling Earth Systems, 14, https://doi.org/10.1029/2021MS002776, 2022.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M.: The ECMWF Operational Ensemble Reanalysis–Analysis System for Ocean and Sea Ice: A Description of the System and Assessment, Ocean Science, 15, 779–808, https://doi.org/10.5194/os-15-779-2019, 2019.