Digital Workshop ESM Initiative 14 June 2021

Preparation and development of the ESM initiative to prepare the natESM project

Agenda ESM Workshop

14 June 2021

09:00 – 09:20	Plenary Welcome and Introduction Jochem Marotzke
09:20 – 09:30	Plenary Statement Bundesministerium für Bildung und Forschung Karsten Hess, Referat 723 – Globaler Wandel; Klimaforschung, BMBF
09:30 – 10:30	Breakout Groups Deliberations on Potential ESM Components
	Clarification of possible remaining questions from the videos
	Strengths and weaknesses of the components
10:30 – 11:00	Break
11:00 – 12:15	Breakout Groups Deliberations on Potential ESM Components (continued)
	Clarification of possible remaining questions from the videos
	Strengths and weaknesses of the components
12:15 – 12:30	Break
12:30 – 13:00	Plenary Feedbacks from the Breakout Groups
13:00	Adjourn

Agenda ESM Workshop

15 June 2021

09:00 – 09:15	Plenary Welcome and Introduction Sarah Jones
9:15 – 10:30	Breakout Groups Assessment of Potential ESM Components
	Compatibility of components with criteria
10:30 – 11:00	Break
11:00 – 11:45	Breakout Groups Assessment of Potential ESM Components (continued)
	Compatibility of components with criteria
11:45 – 12:45	Plenary Feedback from Breakout Groups and Discussions on Potential ESM Components Michael Schulz
12:45 – 13:00	Plenary Open Questions, Wrap-up, Next Steps, and Closure
	Thomas Jung

13:00 Adjourn

Introduction
Purpose of this workshop (1): Overview

- Recap: We originally had 41 suggested contributions of software/model components
- Thanks to all contributors who uploaded recorded presentations; thanks to DLR for hosting the website. Thanks to Mr. Hess for joining us today.
- Today and tomorrow: Clarification/analysis/discussion of contributions, no decisions.
- Discussion will occur in five breakout groups; please self-select where to go. Please make sure a representative is in the "room" where a software is being discussed
  1. Atmosphere, coupler
  2. Land surface
  3. Ocean, cryosphere, biogeochemistry
  4. Infrastructure, data assimilation, diagnostics, workflows
  5. Impacts of climate change
- Breakout groups 2 and 5 still need "supporter"

Introduction
Purpose of this workshop (2): Possible structure

Day 1: Clarifying questions, any topic, such as:

- Purpose of software clear and relevant for ESM?
- Established or potential range of users and applications clear?
- Ownership and institutional support clear?
- Software sufficiently mature to be considered for the development of the strategy?
- Sufficient level of institutional/developing activity?

Day 2: Is the software compatible with the criteria agreed upon at the 2019 workshop, or does it show promise to become compatible with reasonable effort in a reasonable time?

OUTCOME
from the Breakout Groups

BG 1
Atmosphere, Couplers

We discussed the following models and couplers:

- HAMMOZ (Hamburg Aerosol Module – MOZART chemistry model)
- ICON (ICOsahedral Nonhydrostatic Modeling Framework)
- ICON-CLM (DKRZ, DWD, MPI, KIT)
- REMO
- (PALM)
- MESSy Modular Earth Submodel System
- Terrestrial Systems Modeling Platform
- ICON Generalized Interface (GI3) - proposal

BG1
Atmosphere, Couplers

Report day 1: Discussion of Guiding Questions

- Discussion on the issue whether there is already agreement that ICON will be the choice of a future ESM for the atmosphere component: Yes, but the development should be inclusive and incorporate both ICON core group and external user groups
- Subsequent discussion on how ICON core group and user groups could possibly cooperate:
  - D. Klocke reported that the ICON code will be restructured including clear interfaces
  - These interfaces will allow user groups to couple their code to ICON. Timeline will be clear soon.
  - To make this a success this will need engagement between the user groups and the ICON core group before a new structure is set up
- J. Marotzke: The support team that is to be funded by BMBF could possibly support the implementation of the required interface(s)
- The licensing of ICON is still an issue that needs to be resolved; future ESM components should be open sourse.
- No convergence in the group was reached on the issue of how much flexibility a future ESM should finally allow/preserve.

BG1
Atmosphere, Couplers

Results day 2: Discussion of fit to Criteria

- The models show different levels of compatibility with the criteria
- There is agreement that modularity (including well defined interfaces) and performance portability are criteria of prime importance
- Several of the models already significantly invested in getting their compenents exascale-ready
- ICON already mostly fullfills the criteria except for the issue of open source; this is being investigated
- Yesterday we agreed that ICON (here focussing on the atmosphere) should be the atmosphere component model of a future ESM; provided that interfaces for users outside the ICON core-group were provided.

  (Note: ICON core-group will only be responsible for the ICON core-development which includes important components beyond the atmosphere)
- Complementarity between HAMMOZ and ICON-ART, REMO and ICON CLM – potential for collaboration and mutual support was recognized
- We already identified promising/beneficial areas for the „sprints“ of the support group (BMBF proposal)

BG 2

Land, Day 1

Supporter: Barbara Früh, Sönke Zaehle

BG 2
Land, Day 1

- Einführung der Interessen der verschiedenen Gruppen
- Beantwortung der Leitfragen
- Welche technischen Fragen müssen zur Entwicklung einer Stragie beantwortet werden

BG 2
Land, Day 1
Interessen der Teilnehmer

- Beste Techniken
- Einfluss Klima auf Untergrund (und zurück)
- Upscaling
- Kopplungsstrategien
- Landnutzungänderung
- Biogeochemie, Biodiversität
- Anschluss and DestinE
- …

BG 2
Land, Day 1

- Alle Modelle sind valide Entwicklungen; Rolle der Modelle in einer ESM Strategie muss geklärt werden. Wieviel Vielfalt nötig?
- Bedarf besteht bei einigen in der Performance Portability und generischen Kopplungsschnittstelle
- Wie viel Modularität für verschiedene Komponenten und Konfigurationen ist gewünscht/technisch sinnvoll.
- Anwendbarkeit in der universitären und insitutionellenForschung studentischen Ausbildung, und der operationellen Wetter-und Klimavorhersagen
- Skalenkonsistent prozessbaiserte Parametrisierung (Aufhebung der Trennung zwischen Impakt und Prozess-/Feedbackmodellierung)

BG 2
Land, Day 2

- Impakt vs Land-Modell: sollten zusammengeführt werden; Ziel und Zweck sollten bestimmend sein (inhaltlich); wird in Zukunft zusammenrücken; online Diagnostik bis zur vollen Integration
- Akkut: Kopplung mit Atmophäre jetzt Priorität
- Natl. ESM Strat: Nicht nur ein System; auch Modellvielfalt zur Forschung und Verbesserung, die in ESM zurückfliessen können
- Offenheit: Forschungslandschaft kann andocken; Plattform für ESM Forschung, Impakt, Regionaforschung
- One size fits all gibt es nicht (vielleicht nicht in den nächtsen 10 Jahren)

BG 2
Land, Day 2

- Mehrere zentrale Systeme für den jeweiligen (bestimmten) Zweck (nicht mehrere System für den selben Zweck)
- Zentrales System, andockbar mit Modulen anhand von Schnittstellen;
- Ziel ist es Feedbacks besser abzubilden (Mensch, Wald, nicht unbedingt immer Atmosphäre)
- Austausch mit anderen Gruppen; welche Modell(-ketten) existieren; welche Gemeinsamkeiten (kernels) gibt es?
- Rolle Land/Impakt in ESM: Problem der Vielfältigkeit höher als in anderen Bereichen; daraus ergibt sich eine offensive Haltung in ESM wie man integrieren kann/soll -> Sonderrolle

BG 2
Land, Day 2

- Vorläufig: Kernel/Objective Function/Basisfunktionalität
  - Untere atm. Randbedingung (einschl. Mensch)
    - ICON-L(z.Z. JSBACH, TERRA), TSMP(ParFlow-CLM),
  - Wasser- und Energiehaushalt (einschl. Mensch)
    - HD, ICON-L(z.Z. JSBACH, TERRA), TSMP(ParFlow-CLM, CPlantBox)
  - Biogeochemie, Biologie (einschl. Mensch)
    - ICON-L(z.Z. JSBACH4, Quincy), HD, ParFlow(CPlantBox)
  - Feste Erde (einschl. Mensch)
    - Golem-DwarfElephant-Lynx, Fastscape

- Golem: 1, 2, (3), 4, 5(CPUs), 6, 7, 8, 9, 10, 11, 12
- HD: 1, 2, single CPU, single CPU, 5, 6, 9, 10, 11, 12 (Apache)
- ICON-L(JSBACH4): 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12(Quincy)
- ICON-L(TERRA): 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
- TSMP(ParFlow-CLM, CPlantBox): 1, (2), 3, 4, 5, 6, 7(subsurface, land surface), 8, 9, 10, 11, 12

BG3 Ocean, Cryosphere, Marine Biology and Biogeochemistry

Towards a National ESM Strategy

Workshop, June, 14/15, 2021

Facilitators: Sarah Jones and Michael Schulz

BG 3

- Ocean:
  In this section we discussed:
  - the 2 new ocean models being developed at German Institutions: ICON and FESOM,
  - the well-established NEMO model with governance outside Germany but a dedicated user community within Germany that benefit from the connection to the international developments.
  - SCHISM with a strong focus on coastal and estuarine modelling,
  - the wave model WAM.

All models have ongoing development plans with respect to the processes and timescales covered. Further developments are needed to cover a wide range of future challenges including paleoclimate timescales.

Strategies are in place at different levels of maturity for extreme computing and data handling. Collaboration takes place already between the ICON and FESOM developers, with an eye on technical aspects. Potential for intensifying this collaboration should be explored.

BG 3

- Marine biology and biogeochemistry:
  In this section we discussed the models ECOSMO, HAMOCC and RECOM2. These models were initially developed with different perspectives, (e.g. regional or global, carbon cycle or ecosystem) and their subsequent development showed some path dependency.

Land ocean interaction is viewed as a major future challenge internationally. The number of tracers needed in these models present immense challenges for scalability and data handling.

There was interest among the three model developer groups in coming together to consider whether a framework for interactions between biogeochemical modeling groups could create further synergies and be part of the National ESM strategy.

- Cryosphere:
  Sea ice modelling:
  - dedicated sea-ice modelling development takes place at AWI, other developments are outside of Germany. The current Hibler-type models will provide a useful basis in the next decade, but more sophisticated treatments will need to be considered at a later stage
  - The need for consistency in the physical representation of ocean, sea ice, waves and atmosphere ESM components was emphasized.

For the further cryosphere ESM components discussed – Ice-Sheet-Modelling, COSIPY and VILMA - the details of implementation depend to a large extent on the host models. These models have a range of ownerships that are partly outside Germany

General comments:

Some of the criteria being discussed need further clarification: exascale ready, standardization, user friendly.

There was a general concern that recruiting qualified personnel to address the exascale challenge will be difficult.

BG 3

Day 1: Overarching topic: Clarifying questions, any topic

(1) Ocean components

FESOM:
• Timescales on which it can be used? (sub-seasonal to paleo); strong potential for paleo applications; AWI "moves to" paleo applications; challenge associated due to different land-sea masks
• Dyn. Land-Sea Mask? work in progress (focus of future work)
• Multiscale Approach: previous issues regarding mixing solved? community effort; horizontal mixing can be improved by hi-res in specific areas
• Numerical efficiency?: approx. same as NEMO
• FESOM investments: computing, ESM longer timescales

ICON:
• how widespread is user community? Is there a community beyond core developers? Main base currently "in-house" (DWD, MPI-M, KIT, DKRZ); COSMO Consortium, Hereon (on biogeochem) and C2SM in Switzerland; user group is increasing with focus on hi-res modeling (<= 1 km) and decadal prediction
• re biogeochem.: "innovative" approaches for efficient use of "slow" biogeochem. component (incl. pollution)
• ICON-SEAMLESS concept for NWP-seasonal-decadal and beyond;
• Bundeswehr funding for coupled atmosphere – ocean at weather timescales.
• Dyn Land Sea? not tackled yet (importance recognized)
• Paleo Appl.? Thousands of years possible; further development through collabs envisioned

BG 3

Day 1: Overarching topic: Clarifying questions, any topic

(1) Ocean components

NEMO

Community model -- is it a tool or what are the development contributions from GEOMAR? - here as a group of interested users. Developing configurations (blue water, coastal and regional)

- special expertise in D: wave-ocean coupling; parameterizations for wet-dry conditions; implementation/testing of couplers
- Consortium: need to supply a dedicated set of developers, grouping sub critical to enter the consortium -- if group enters the consortium, more control. BSH thinking about this. 1-2 FTEs; need to join to have direct influence on decision making and how this will develop in the future.
- (need to get back to what is missing in other models compared to NEMO --, also the question of the fit-for-future-HPC etc.)
- gain from "another" Nemo-based ESM? NEMO intended as an option (not as core) in order to stay connected at the European Level; risk of getting disconnected from international developments that could create synergies
- NEMO linking to services --- needed as tool, working with set-ups
- zoom-in tool very beneficial for regional applications

BG 3

Day 1: Overarching topic: Clarifying questions, any topic

(1) Ocean components

SCHISM

- designed and widely to capture processes in land-ocean transition zone
- stakeholder-driven model development

WAVE

- commonality with ECWMF? started from same core; now different branches (3: ECWMF; Meteo-France, Hereon) --> only Hereon branch is open source

HGF strategy w.r.t. ocean model?

- ongoing process;
- finding the right balance between large developments and established models (incl. collaborations)
- Helmholtz centers haven't agreed on one ocean model for future --- further foster collaboration --
- so National ESM challenge is very much a Helmholtz Challenge

BG 3

Day 1: Overarching topic: Clarifying questions, any topic

(1) Ocean components: **Fitness for future IT architecture / data handling?**

FESOM: exascale/data --> strong commitment ("fast track" mode of operation not entirely sustainable as yet); collab. at European level (PyFESOM as an initial step)
bottleneck: finding qualified personnel, Question of becoming fit for future HPC is not whether but when
ICON: exascale readiness is core priority in good progress ("fast track" mode of operation)
big data: becoming more important; development bottom-up driven
challenge: "out-of the box thinkers“

NEMO: Exascale: active group at European level (apparently mainly in UK), DSL&GPU; use community has large-scale high res applications big data: challenges are mainly seen in the workflow; AI initiative in HGF (appears currently only loosely connected)
SCHISM: big data: specific challenges in coastal regions w.r.t. to model validation using different streams of obs.
WAVE: exascale readiness tackled in coop. with European partners; has to be considered in the context of the ocean model to which wave model is coupled

Observation from Carsen Lemmert: *One observation on the ocean models (maybe for later discussion). We seem to be building separate „ecosystems“ of models (FESOM, ICON, SCHISM, NEMO), but less to think about synergies by explicit interfaces between the systems and modularisation of components of the systems for interchange.*

BG 3

Day 1: Overarching topic: Clarifying questions, any topic

(2) Marine biology and biogeochemistry

ECOSMO
- global applications? in the process; not validated yet
- specific modules (inorganic sediments.; cyanobacteria....) developed based on regional needs
- not used in global ESM or paleo context (multidecadal as yet)

HAMOCC (ICON)
- which version? most recent (all HAMOCC development now part of ICON)
- possibility of concurrency: allows for hi res in biogeochem.
- coupling with sediments and land-ocean transition
- model years per day in low res version (HAMOCC/ICON)? @160 km 200 yrs/day
- scaling between no. of tracers and performance: scaling is non-linear

REcoM2
- efficiency: ongoing work to enhance the scalability w.r.t. to large number of tracers (not all tracers require flux correction)

General Questions:
- do the model components being discussed here cover the range of processes needed for current and future applications
- Land-Ocean interaction not a strength of German community

BG 3

Day 1: Overarching topic: Clarifying questions, any topic

(2) Marine biology and biogeochemistry: the future HPC/ Big data question

ECOSMO:
- future gaps: pollution and interaction with ecosystem; ecosystem: size spectra; “Lagrangian” organisms; coupling to land component (morphodynamics); integrating higher trophic levels; link to sea-ice biota/biogeochem.
- future HPC challenges: flexibility of the coupling --> dependence on "host model" (ocean) should be minimal in order to use different hydrodynamics

HAMOCC:
- future gaps: coupling with other ESM components on timescales up to centuries
- future HPC challenges: immense data output; analysis needs to be part of the experiments (adaptive output)

REcoM2:
- future gaps: making the biological pump more robust; land-ocean coupling w.r.t nutrients, carbon
- future HPC challenges: dependence of ocean model development

Day 1: Overarching topic: Clarifying questions, any topic

SPONSORED BY THE
Federal Ministry of Education and Research

(2) Cryosphere (incl. sea ice) – Sea ice components in ocean models

FESOM: finite element model; AWI committed to development; current approach scientific valid for down to ~1km scale; computational demand relatively small
General remarks:
main developments outside Germany
challenges: beyond Hibler-type models; inclusion of biogeochemistry
community effort in sea-ice models (Los Alamos) that is supported from Germany (AWI)
Consistency between components: ice, ocean, wave, atmosphere important!

COSIPY

./.

IceSheetModeling (AWI)

- Sophisticated surface energy mass balance? "semi complex EBM"; option to deal with firn
- versatile coupling options
- open source; community effort
  → HGF strategy: ISSM will be part of AWI ESM

VILMA

requirements from ESM for its use? coupling is rather flexible (remapping of surface mapping to Legendre grid); should be possible to be used with ICON
limiting factor: VILMA: spectral FE model (very efficient); other models use more expensive FE approaches

BG 3

Day 1: Overarching topic: Clarifying questions, any topic

(3) Future challenges

- Ice-sheet models
• coping with melt at surface of ice sheets (requires hydrological models instead of EBM only)
• clear interface between ice sheet and atmosphere
• representation of sliding processes
• Gap between ice and climate modeling communities (e.g. is firn part of ice or atm)

- Solid Earth models
• decreasing resolution to km scale
• thermodynamic aspects (viscosity and geothermal heat flux)

- Snow Model
water transport; lack of permafrost
inclusion of ice-biogeochemistry

BG 3

Day 2: Overarching topic: Software compability

- General:
  Definition of “exascale ready” and how this applies to the different components --> needs to be defined by applications rather than HPC perspective (i.e., what is required to run specific configurations with typical timescales) considering scalability, heterogeneous architecture, ...

Following criteria are unclear:
- standardization
User friendliness from university perspective --- can it be used by a master's student? ---- was discussed at length in DFG roundtable, information on what is needed to use model in teaching framework --- need to follow up on that.

NB: users don't care about scalability but about efficiency and it could be interesting to compare efficiency between the different ocean models in this group, esp. ICON, FESOM, NEMO. (potential to carry out MIP for typical configurations, e.g. NWP and paleo); existing comparison: NEMO and FESOM comparable for similar hi-res setups
Comment: T. Kleiner*: Long integration times will be even more relevant with interactive ice sheets. For me this raises the question of how to initialize a fully coupled ESM.*

Comments from Chat:

@André: how do you assess efficiency if you cannot run a simulation on one core?
Most of our applications are too large to measure with one core, hence you cannot define speedup and with that efficiency cannot be computed. Still you can measure scalability well. Or to put it in other words: if you compute small problems, you can measure speedup and efficiency, but than you have so few DOF/core, that a code cannot scale well Don’t we need another definition of metrics here?
@Anegelika: Maybe one could define a benchmark ocean simulation with a given number of degrees of freedom on a comparable cluster with a given number of cores. However, as Sergey pointed out, carefully defining and carrying out a benchmark is quite some work.
@André: yes, such benchmark problems are a great way to go. Ice sheet modellers are running benchmarks (ISMIP6), that would be really good basis for such scalability studies and yes, you would need to define targeted DOF to allow comparison between the different codes. This is really the way forward from my perspective

BG 3

Day 2: Overarching topic: Software compability

- **Ocean Components**

FESOM:
already working in comprehensive ESM context
ESM tools --> workflow manager
global to local OK by design; in addition, dedicated coastal version in prep.
exascale readiness: initial steps using ACC ("fast track") (goal 10e5 GPUs)
scalability of workflows: additional effort required
portability: OK
modularity: in progress (major steps expected within two years)
DA: possible with FESOM as well as in coupled system
Synergies with ICON code --- moving parts between components?
Diagnostic -- connection to ESMVal
User friendly: Documentation available; version control via github
open source license (GPL)
full ownership (AWI)

Q: use for commercial purposes without restrictions? yes
Q: incorp. of community in development? generally open; most input at this point from AWI / assurance of quality control ?

BG 3

Day 2: Overarching topic: Software compability

- Ocean Components

ICON:
already working in comprehensive ESM context
global to local OK (grid refinement; nesting)
exascale readiness: preparations in progress (e.g. JP HPC system); performance of code not the issue; I/O is major challenge; [current model is faster than I/O]
scalability of workflows: well established with version control and gatekeepers for quality control
portability: in progress
modularity: OK (people are able to setup model in short amount of time)
DA: PDAF for ocean in development;
diagnostics: Python based; exchange with FESOM group in TRR181 (Energy transfers)
user friendly: yes
documentation: can be improved; currently too distributed
Version control: established
open source: working towards this, need to define what we mean by open source, open issue due to status of DWD (constraints due to DWD Gesetz and BHO) ; currently: complete availability to research community
ownership: the four ICON institutions

Q: DA w.r.t. coupled system? open discussion in seamless project, which only started this year
NB: good collaboration between FESOM and ICON-O (incl. sea ice) groups (partly through TRR181) as well as exascale readiness
Q: Tendency to converge? not within few years. Some convergence w.r.t. infrastructure as first steps possible; question of convergence could be re-evaluated in a few years; keeping options open seems best strategy at this point
ICON and FESOM: collaboration between developers, exchange on configurations, diagnostics, potential convergence on infrastructure; longer term need to be self-critical about critical nass for development. Work together on infrastructure part. Not in a position to choose between models in next 2 years or so.

Day 2: Overarching topic: Software compability

- Ocean Components

NEMO:
typically in ocean-only and climate modes
global to local: OK (2-way nesting well established)
exascale readiness: identified need (HPC working group), not as far as FESOM and ICON
scalable workflows: seems OK
portability: given (wide range of user)
modularity: OK
DA: exists but doesn't not come with code
diagnostic capability: online does not come with code
documentation: exists; very important are test case
traceability: version control exist
open source: CeCILL (French version) of GNU license
performance expected to be significantly improved (as part of EU project)
consortium membership requires "critical mass" in Germany
Q: potential for synergies with FESOM/ICON communities? comparison of standard configurations (MIP-type)
NEMO/ICON/FESOM: set up / compare standard configurations and compare to community; understand if not selected as not German development, but there is a German community using NEMO and that should be taken into account. Part of European activities including COPERNICUS.

BG 3

Day 2: Overarching topic: Software compability

- Marine biology and biogeochemistry

SCHISM

interfaces with focus on processes that are relevant for coastal areas
no global setup; model was designed for regional to local scales, unstructured grid, focus for estuary
exascale ready? "long way to go"
scalability: scales well for up to 1000 cores
portability: runs on more than 30 HPC systems
modularity: yes
DA: exists
diagnostics: ????
user friendly: documentation, wiki, large user community
traceability: version control exists
open source: GPL license
ownership: US but Hereon closely connected to developers ("influence")
Q: traceable German contributions?: interfaces between modules; parameterizations; Hereon has copyright on certain code parts, provide cases and test them,

BG 3

Day 2: Overarching topic: Software compability

- Marine biology and biogeochemistry
- WAVE:
  interfaces are flexible; has been coupled to different ocean and atm. models (e.g. OASIS); to be integrated in ICON
  global to local possible
  exascale readiness: no concrete steps yet
  scalable: given
  portability: runs on many HPC systems
  DA: exists (global and regional level)
  Diagnostics: OK
  user friendly; documentation exists; test cases
  version control: exists
  open source: owned by Hereon

  General: different functionalities and capabilities - need to discuss that as well as technical criteria
  Q: Framework for interactions between biogeochemical modeling groups? Not yet ---- could be an outcome of this strategy (NB: groups are in different topics in POF). Started in Helmholtz ESM activity -- but not intense
  Recognize need/advantage to get together and discuss this ---- will try and do this before the next workshop
  ECOSMO started from ecosystem whereas HAMOCC/REcoM originated from C-Cycle -->perspectives for synergies before fall workshop seems possible [would be a great outcome] could also involve work by Iris Kriest @ GEOMAR

BG 3

Day 2: Overarching topic: Software compability

- Marine biology and biogeochemistry

ECOSMO

interfaces: yes w.r.t. to carbon; used with several ocean models
global to local: yes; parameter might need adaptation
exascale: related to host model
scalable workflows: related to host model
portability: related to host model
modularity: given
DA: related to host model; only rudimentary for biol. part
diagnostics: yes
user friendliness: work in progress (documentation)
traceability/version control: given
open source/ownership: open but linked to license to FASOM (SPELLING???) unclear which license; Hereon but also Norway

HAMOCC (ICON)

generally as ICON; some older parts of code still need conversion

BG 3

Day 2: Overarching topic: Software compability

- Marine biology and biogeochemistry

REcoM2

generally as FESOM (only differences noted)
interfaces: called directly from FESOM; no specific cpl.
global to local: yes
scalable workflows: tests w/ FESOM --> seems to scale well
modularity: pt. of ESMTools
DA: PDAF for biol. and some BGC variables
user friendliness: draft of documentation
traceability/version control: git repositories
open source/ownership: open and AWI is owner

BG 3

Day 2: Overarching topic: Software compability

- **Cryosphere**

Challenge to have a modelling system going to km scale on one hand and to1000s of years on other hand

COSIPY

interfaces: works w/ WRF and COSMO
global to local: yes but generally applied locally
exascale: ./.
scalable workflows: ./.
portability: yes
modularity: yes
DA: exits
diagnostics: user friendly (Py)
user friendliness: user network
traceability/version control: github, automatic testing for new code
open source/ownership: open, Uni Erlangen

BG 3

Day 2: Overarching topic: Software compability

- Cryosphere

IceSheetModeling (AWI)

interfaces: with atm and ocean not with VILMA
global to local: in principle yes
exascale: answer needs definition.....comments on efficiency, not GPU
scalable workflows
portability: yes
modularity: yes
DA: yes but via complete adjoint
diagnostics: yes (also MATLAB and Py interfaces)
user friendliness: documentation to be improved
traceability/version control: yes
open source/ownership: community effort

BG 3

Day 2: Overarching topic: Software compability

- Cryosphere

VILMA

interfaces: via netcdf
global to local: yes (to regional)
exascale: not relevant
scalable workflows: yes
portability: proven
modularity: is itself a module
DA: not applicable
diagnostics: nothing specific
user friendliness: documentation in prep.
traceability/version control: yes
open source/ownership: open; license currently discussed at GFZ, who is owner

BG4 Infrastructure, Data Assimilation, Diagnostics, Workflows

Towards a National ESM Strategy

Workshop, June, 14/15, 2021

Facilitators: Hauke Schmidt, Joachim Biercamp

DKK Deutsches Klima Konsortium
SPONSORED BY THE Federal Ministry of Education and Research

Software to be discussed

	Tool	Institution	People	Purpose
Data Assimilation	DACE (Data Assimilation Coding Environment)	DWD	Roland Potthast	Data Assimilation
	PDAF (Parallel Data Assimilation Framework)	AWI	Lars Nerger	Data Assimilation
Coupling, model/experiment configuration	DKRZsw	DKRZ	Hendryk Bockelmann, Panagiotis Adamidis	Other (coupling, communication, IO)
	ESMTools	AWI	Dirk Barbi	Workflow (model/experiment configuration)
	Xarray-simlab	GFZ	Jean Braun	Workflow (model/experiment configuration)
Diagnostics and Evaluation	ESMValTool (Earth System Model Evaluation Tool)	DLR	Veronika Eyring, Axel Lauer, Birgit Hassler	Other (Postprocessing, Model Evaluation)
	HeAT	FzJ	Klaus Görgen	Diagnostics
	SAD wavelets	Uni Bonn	Petra Friederichs	Diagnostics

BG4 Infrastructure, Data Assimilation, Diagnostics, Workflows - notes (1)

- With respect to data assimilation, coupling, model/experiment configuration, diagnostics and evaluation, a lot of comprehensive tools, useful for a national ESM strategy have already been developed and are used in German modelling efforts
- We have identified complementarity and potentials for synergy of several approaches
- Relevant criteria are in general fulfilled unless mentioned otherwise below

- Data assimilation:

    - DACE:
        - comprehensive DA system for NWP, large developer basis, comprehensive selection of atmospheric observation operators
        - Exascale-readiness is work in progress; generic ICON issue for open source; complex system that is in parts difficult to change for users
    - PDAF:
        - originates from ocean applications but is flexible to use, “model agnostic”
        - Exascale-readiness (porting to GPUs) hasn’t started (but use of gpu versions of blas/lapack can be used)
    - Complementarity of the two approaches; DACE is necessary for comprehensive atmospheric DA, PDAF is necessary for flexible DA in a variety of components an applications; there are examples for the application of both tools in one ESM
    - We feel that jointly the two tools cover a wide range of DA necessities; however model adjoints (4D-Var) are currently not covered

DKK Deutsches Klima Konsortium
SPONSORED BY THE Federal Ministry of Education and Research

- Coupling, Model/experiment configuration

    - DKRZsw:
        - coupler (YAC); inter-process MPI communication (YAXT); parallel output (CDI-PIO) to enable HPC usability of ESMs;
        - mostly applied for models heavily in use at DKRZ machine (MPI-ESM, ICON)
        - exascale-readiness is what the tools are aiming at; end-user documentation can be improved
    - ESMTools:
        - Framework for downloading, compiling, running and organizing ESM code;
        - Standardizing workflow in a user-friendly way is the main purpose of the tool
    - Xarray-Simlab:
        - Python-based coupler of component models
        - Well-suited for exploratory modelling and education purposes, in principal open to any code that could be called from python, but it would be challenging to couple existing complex components; targeted for interactive applications and analysis-ready IO
    - The three tools are complementary, having different targets. Specific users may have workflow requirements that are not yet covered by the tools mentioned here.

BG4 Infrastructure, Data Assimilation, Diagnostics, Workflows - notes (3)

- Diagnostics and Evaluation:
    - ESMValTool:
        - diagnostic tool for routine evaluation of Earth system models in CMIP
        - Currently requires CMOR output; applicability to native outputs is being developed
        - Exabyte-readiness (very large or very high resolution ensembles) is being worked on
    - HeAT:
        - Toolkit for NumPy-like big-data analytics and processing (including ML)
        - Supports heterogeneous HPC systems and different architectures
        - Relatively new development, but strong institutional support
    - SAD wavelets:
        - Tool for model diagnostics beyond point-wise metrics (spatial patterns)
        - Several criteria not yet fulfilled
        - Offer to enhance existing tools with a new method
    - Synergies:
        - Discussion have started between ESMValTool and HeAT representatives on the possible use of HeAT within the ESMValTool framework
        - SAD could be integrated e.g. into ESMValTools if considered useful

BG5 Impact

Towards a National ESM Strategy

Workshop, June, 14/15, 2021

Facilitators: Sabine Attinger

BG5 Impact

Presentation of a broad portfolio of „impact models“ ranging from hydrology, forest / grass / agricultural ecosystems and urban environments

Typical variables

- hydrological state variable + fluxes, biomass production, crop yield, nutrient fluxes
- nitrogen + carbon fluxes at the landsurface
- sediment transport, urban climate
- and very typical management → land-water

→important for topics heavily discussed nowadays

- Landtransformation
- Agrarwende
- Water strategy of government
- Biodiversity

←→address different stakeholders (not only DWD) UBA, ministries, water suppliers…

BG5 Impact

Overview over the very professional code development and the groups behind it and several have already their code complete to ESM models

→ we are ready for this initiative

Issues to be discussed:
- modularity
- technical implementation of interfaces and couples

BUT also
- there are conceptual design differences in these different model families → to be discussed before coupling between impact models and ESM models
New opportunities by new global data sets

Modular → model-chains
One way ←→two way
container solution for users

Overwhelming capacities / capability of the existing codes

- Each code has its own strength / main focus → combining modular parts of our mode to build model chains from weather / climate → impact

- The specific "impact"s stakeholder driven, question driven!
- In most of the case: one-way coupling but there is also two-way coupling in particular for the vegetation models → land schemes