CRC1502
DETECT
https://sfb1502.de/

Status of TSMP2 by DETECT/Z04

2025-05-14 I Stefan Poll[1,2,3], Marco van Hulten[4], Johannes Keller[1,2,°], Paul Rigor[1,2,°],
Daniel Caviedes-Voullième[2,3,°], Klaus Goergen[1,2] (Z04 and °affiliated)
Under the auspices and with support of groups from Stefan Kollet[1,2] and Harrie-Jan Hendricks-Franssen[1,2]
Supported by Kaveh Patakchi Yousefi[1,2], Niklas Wager[1,2,*], Carl Hartick[1,2,*], Muhammad Fahad[1,2],
Ana Gonzalez-Nicolas[2,3]

[1]Institute of Bio- and Geosciences (IBG-3, Agrosphere), Research Centre Jülich (FZJ), Jülich, Germany

[2]Centre for High-Performance Scientific Computing in Terrestrial Systems (HPSC TerrSys), Geoverbund ABC/J, Jülich, Germany

[3]Simulation and Data Lab Terrestrial Systems (SDLTS), Jülich Supercomputing Centre (JSC), Research Centre Jülich (FZJ), Jülich, Germany

[4]High-Performance Computing and Analytics Lab (HPC/A-Lab), University of Bonn, Bonn, Germany

*New affiliation

Goals

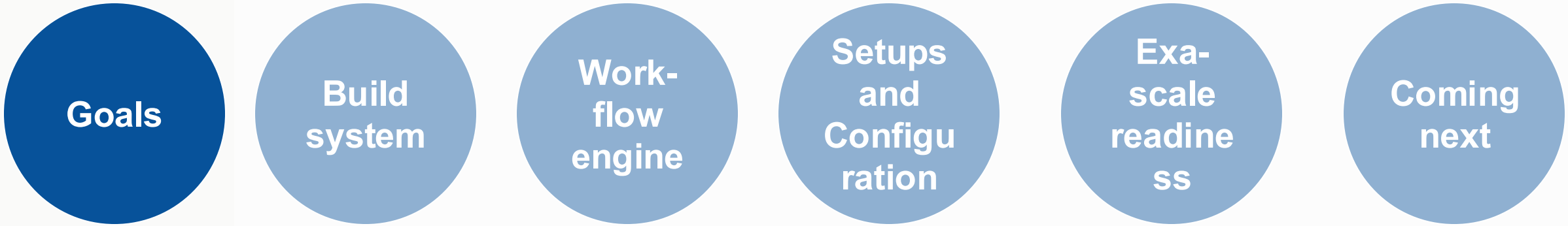
Build system

Work-flow engine

Setups and Configu ration

Exa-scale readine ss

Coming next

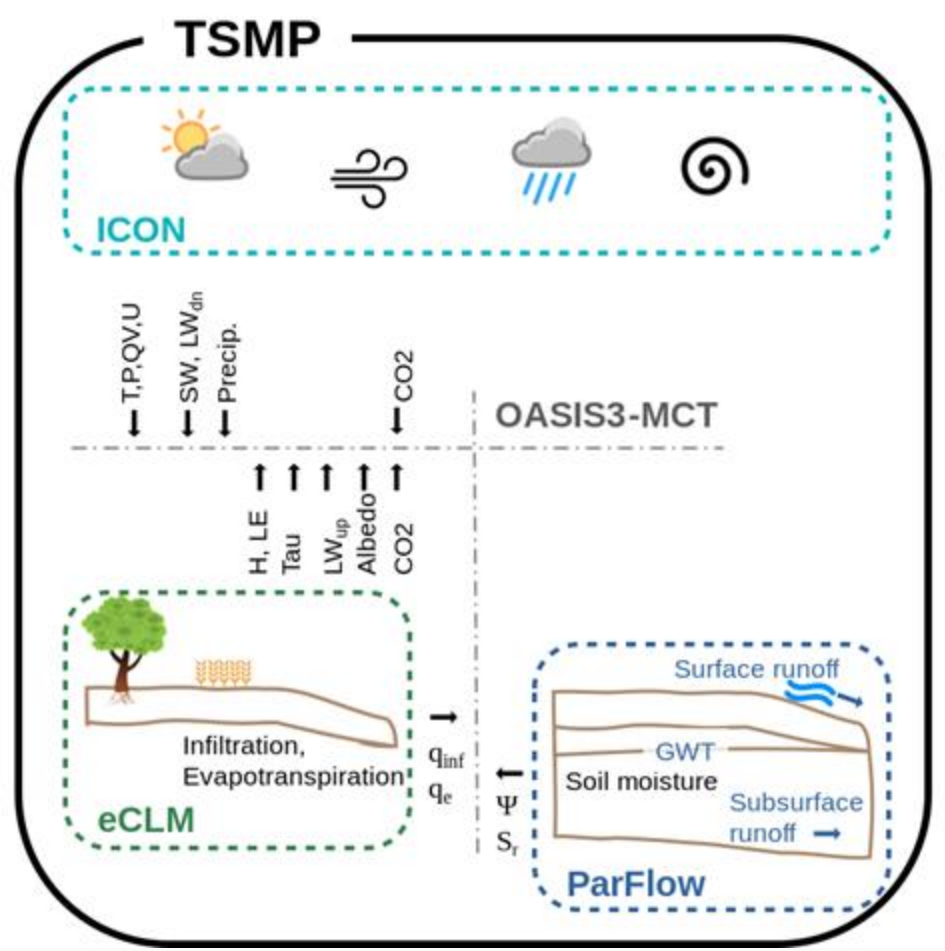


Goals and Overview

TSMP overview, coupling, goals in DETECT I

TSMP2 framework – Overview

Terrestrial Systems Modelling Platform v2 (TSMP2)



- Modular coupling design (MPMD + HC)
- Open source (framework + components)

ICON – v2024.07

Zängl et al., 2015

Atmospheric model

eCLM – Fork of NCAR's CLM5.0

based on Lawrence et al., 2019

Land-surface model

ParFlow – v3.13

Kollet and Maxwell, 2006; Kuffour et al. 2020

Integrated hydrological model

OASIS3-MCT – v5.0

Craig et al., 2017

External coupler

TSMP framework – Atmosphere-Land coupling

Explicit coupling approaches after Polcher et al. (1998)

TSMP1

Exchange coefficients approach

- Coupling via exchange coefficients ($C_{\theta,q,M}$) and surface variables
- Calculation of surface fluxes

TSMP2

Fixed fluxes approach

- Coupling via turbulent fluxes
- Direct usage of surface fluxes

Surface flux calculation in ICON

$$H_0 = -\rho\, c_p\, C_\theta\, |v_h|\, \Delta T$$

$$(L_v E)_0 = -\rho\, L_v\, C_q |v_h|\, \Delta q_v$$

$$M_{i,0} = -\rho\, C_M\, |v_h|\, \Delta u_i$$

TSMP2 framework – Land-Subsurface coupling

Explicit coupling approach; coupling via Richards equation

Surface and subsurface component models share the same vertical discretization

- Infiltration and root zone transpiration (eCLM) show up as Q
- Pressure head (ParFlow) replaces soil matric potential (eCLM)
- ParFlow overrides soil hydraulic properties specified in eCLM
- Soil water is computed to ParFlow and then passed to eCLM

Richards equation:

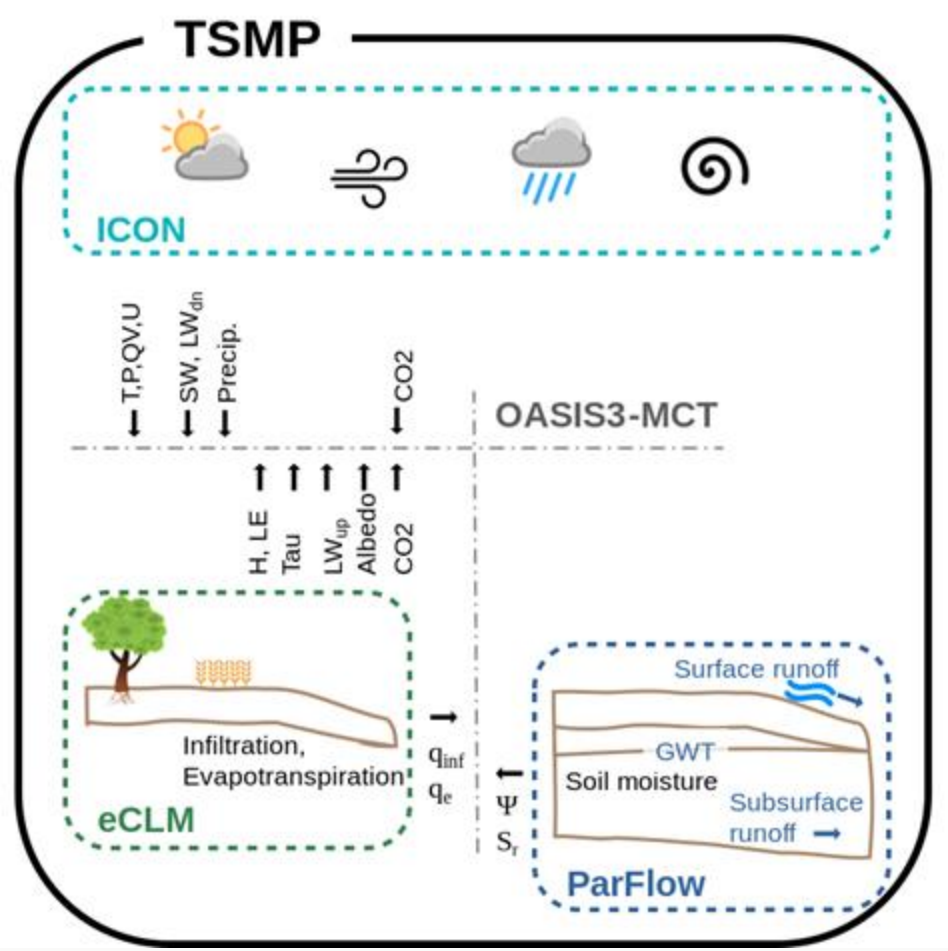
$$S(p)S_s\frac{\partial p}{\partial t} - \frac{\partial(S(p)\rho(p)\phi)}{\partial t} - \nabla\cdot(\mathbf{K}(p)\rho(p)(\nabla p - \rho(p)\vec{g})) = Q$$

where:

p – pressure head
S – water saturation
Φ – porosity
S_S – specific storage coefficient
$\mathbf{K}(p)$ – hydraulic conductivity tensor
ρ – water density
g – gravity

TSMP2 framework – Context and overall usage

Terrestrial Systems Modelling Platform v2 (TSMP2) in DETECT



- The Integrated Modelling System (IMS) is important for hypothesis testing in DETECT CRC
- IMS can simulate feedbacks and interactions in groundwater-land-atmosphere system

- Fully coupled TSMP2 is usable, all components and software tools in the workflow are ready
- Increase user friendliness
  - Front end in the building system
  - Introducing a runtime environment

- TSMP2 component models in use by DETECT WPs; the fully coupled is not used yet, though there is potential
- Fully coupled TSMP1 (COSMO-CLM3.5-ParFlow) is “workhorse” for hypothesis testing, e.g. D02

TSMP2 fully coupled simulation experiments, towards RESM

Goal: Must have TSMP2 fully usable for DETECT / Phase 2 and published beforehand

Workflow is ready: simulation setup (HPC system, software environment, built system, domain, external parameter fields), model configuration, compute configuration, initial conditions, automized workflow engine (BC pre-processing, run-control / -monitoring, (post-)processing, archiving), dissemination (full data lifecycle)

Planning – once data is produced more analyses are possible, Z04 will conduct the TSMP2 3km baseline run:

1. 2025/Q1-Q2: TSMP2 introduction paper (e.g., GMD(D)), using 12km and idealized tests; proof of concept; baseline paper for TSMP2
2. 2025/Q2-Q3: TSMP2 high resolution, coupled, pan-EU frontier experiment presentation and evaluation paper (e.g., GRL), short test / annual cycles / individual weather events / multi-year eval runs

Baseline setup: 1hr ERA5-driven, direct nest, CLMcom ICON configs, planned/needed: 2000-2020 (time-slices), transient GHGs and aerosols, static land cover, no HWU, needs IC from ParFlow standalone run EUR-0275, run “in the background”

Goals | **Build system** | Work-flow engine | Setups and Configuration | Exa-scale readiness | Coming next

Build System

Developments in build system, component models and beyond

Building within TSMP2 framework

New TSMP2 frontend for more convenient building of model components

Build your model in just two steps:

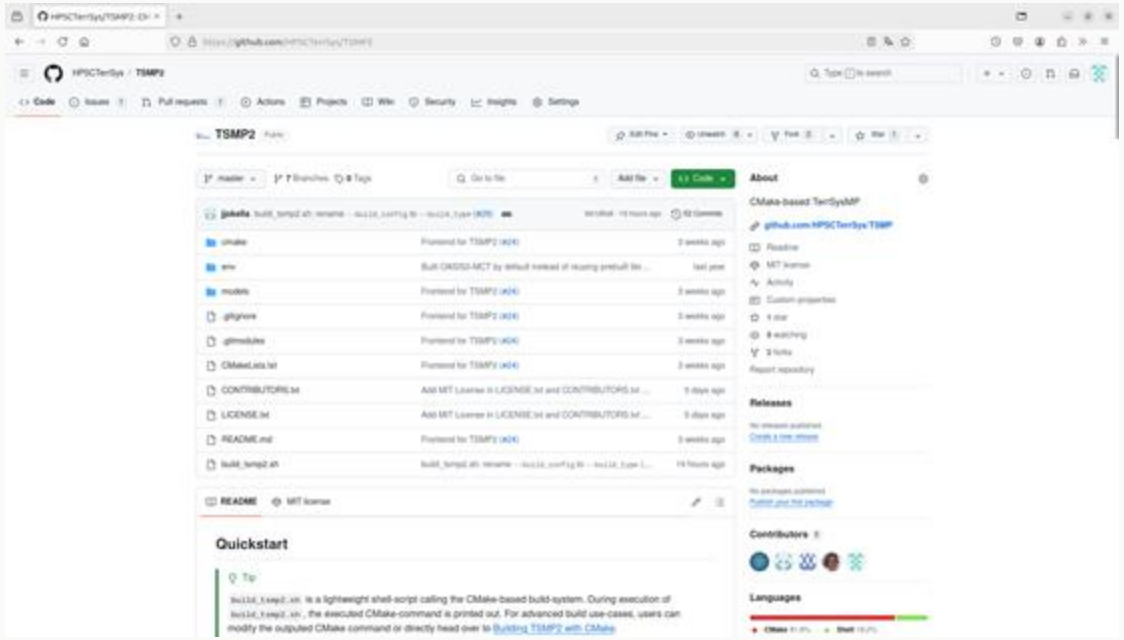
1) Clone repository: git clone https://github.com/HPSCTerrSys/TSMP2
2) Build model components: ./build_tsmp2.sh --ICON --eCLM --ParFlow

Supported machines (by default):

- JUWELS, JURECA-DC, JUSUF
- JEDI / JUPITER
- MARVIN
- Ubuntu

Multiple Environments possible

- Intel, GNU



Model improvements & upgrades

Optimisation of the coupling

Update of component models

i) ICON: update to open-source release, synced with icon/icon in background

ii) eCLM: update procedure in planning

iii) ParFlow: follows upstream

ICON-eCLM coupling:

i) eCLM first receive from ICON, then read streams → better flow of data

ii) Do not couple lateral boundary points (halo-region)

TSMP2-PDAF: Data assimilation with TSMP2

Collaborative development environment

Build

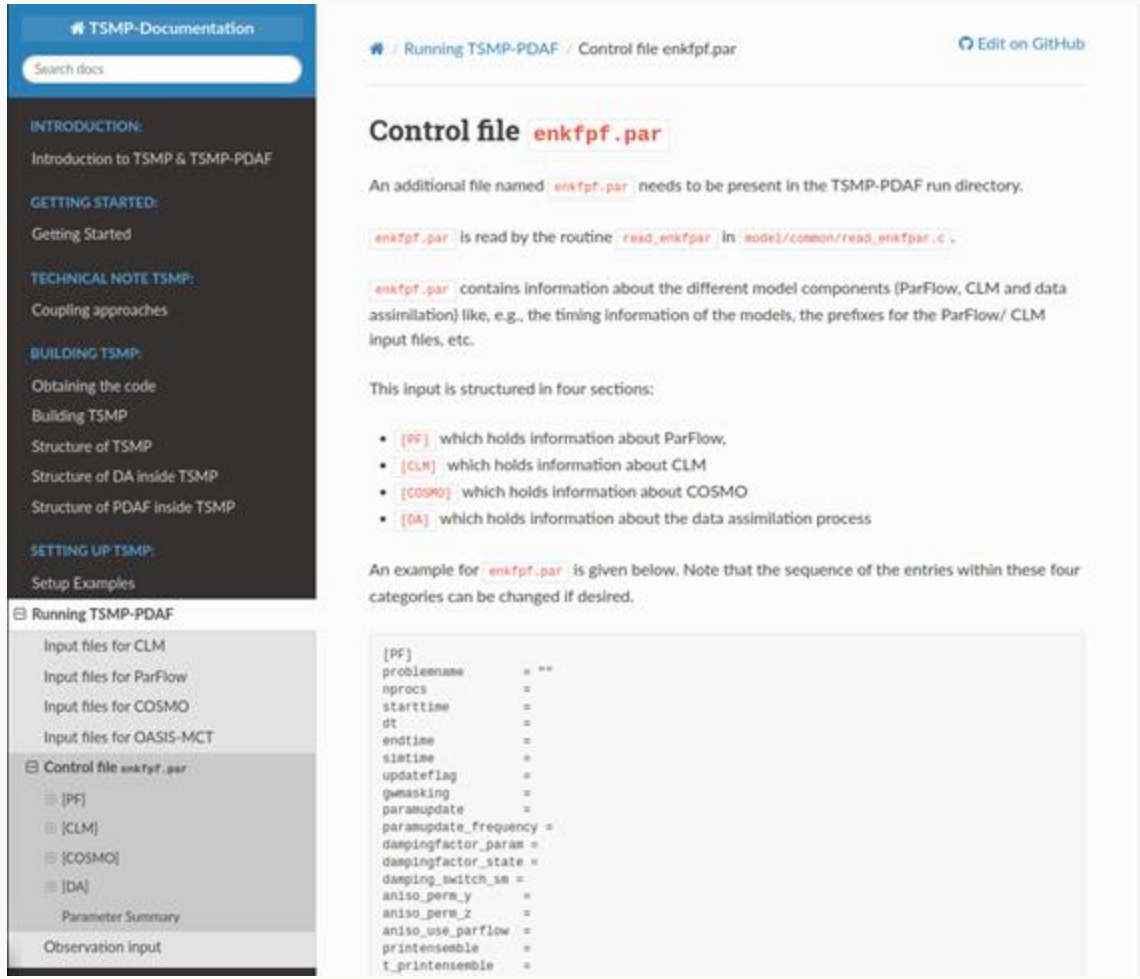
- TSMP2 repository: https://github.com/HPSCTerrSys/TSMP2/
- New frontend for simplified user experience: ./build_tsmp2.sh --eCLM **--PDAF**

Documentation

- https://hpscterrsys.github.io/TSMP/content/setup_tsmp/input_enkfpf.html

Issues

- In-repo (TSMP2, eCLM, PDAF,...)
- DETECT (https://gitlab.jsc.fz-juelich.de/detect/detect_z03_z04/discussions/-/issues)



TSMP2-PDAF: Data assimilation with TSMP2

Status and Plans

Status

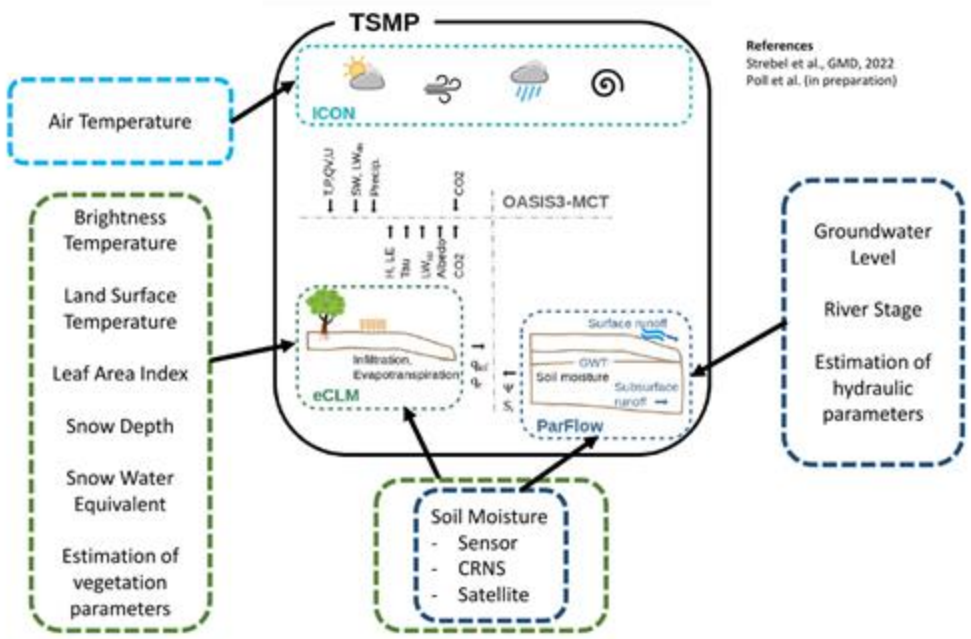
- Recent developments:
  - New observation types in TSMP-PDAF: Snow-Depth / SWE, LAI (point-scale), LST
  - Local filters (e.g. LESTKF) for eCLM-PDAF

Plans for 1st phase of DETECT

- OMI-interface, GRACE-DA currently in University Bonn branch (Yorck), include into TSMP-PDAF
- Fully functional (assimilation into point-scale and distributed simulations)
  - LST-DA
  - Snow-DA (SWE / Snow-Depth)
  - LAI-DA

Proposals for 2nd phase of DETECT (new in bold)

- **Flexible multivariate assimilations** of all observation types, i.e. SM, in-situ ET, GRACE data, LST, Snow variables, LAI
- **Hypothesis testing** with TSMP-PDAF
- **Snow-cover DA**



Goals | Build system | **Work-flow engine** | Setups and Configuration | Exa-scale readiness | Coming next

Runtime environment – TSMP2 workflow engine

Convenient way to perform simulations

Why using TSMP workflow engine?

- Simulation control and reproducibility
- Pre-processing, simulation, post-processing, visualization

Features

- Flexible usage of any model component and combination in TSMP2 realm
- Usable for component models standalone and coupled systems
- For test, production and sensitivity runs
- Download and “direct” start

Technicalities

- Version controlled (git based)
- Git repository at github
- Shell-based / simple structure
- Choice of experiment based on submodules

Workflow engine as the skeleton

- Is not functional in itself
- “Knows” where to set the parts

Git submodules make it complete

- for model source code, model namelist
- for static fields
- for post/pre-processing scripts
- ….

TSMP2 workflow engine

Structure of TSMP2 WFE

TSMP2_wfe/

|---- ctl -> managing the simulation

|---- nml -> model namelists

|---- run -> number crunching

|---- dta -> data

|---- src -> source codes

Two configure files to manage configuration / setup

|---- master.conf -> general settings

|---- expid.conf -> model specifications

master.conf (excerpt)

```
# main settings
MODEL_ID=ICON-eCLM-ParFlow #ParFlow #ICON-eCLM #ICON-eCLM-ParFlow #ICON
EXP_ID="eur-11u"
CASE_ID="" # identifier for cases
conf_file="" # configure file for cases, also possible ${EXP_ID}.conf

# main switches (PREprocessing, SIMulations, POSt-processing, VISualisation)
lpre=( false false false ) # config, run, cleanup
lsim=( true true true ) # config, run, cleanup
lpos=( false false false ) # config, run, cleanup
lvis=( false false false ) # config, run, cleanup

# time information
cpltsp_atmsfc=900 # coupling time step, atm-sfc, eCLM timestep [sec]
cpltsp_sfcss=900 # coupling time step, sfc-ss, ParFlow timestep [sec]
simlength="1 day" #"23 hours"
startdate="2017-07-01T00:00Z" # ISO norm 8601
inidate=${startdate} # ISO norm 8601, fix for entire simulation
numsimstep=1 # number of simulation steps, simulation period = numsimstep * simlength
```

Goals

Build system

Work-flow engine

Setups and Configuration

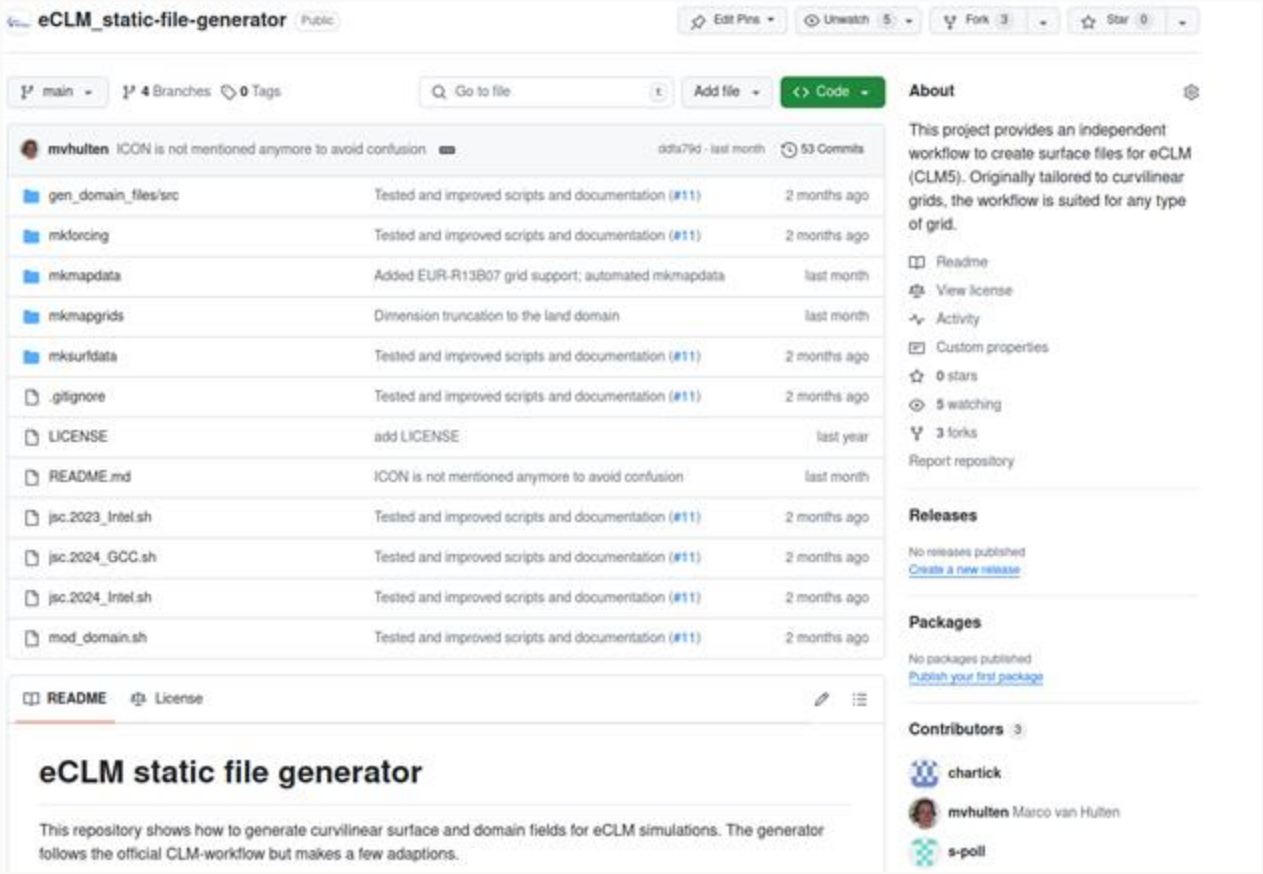
Exa-scale readiness

Coming next

Generate of new domains

eCLM static file generator

- Creation of new static parameters
- Make consistent static fields across TSMP2 compartments
- Creation of eCLM/CLM5 atmospheric forcing files



TSMP2 fully coupled simulation experiments, 3km frontier runs

Model configurations, 12km (R13B05, EUR-11), 3km (R13B07, EUR-0275)

- Climate mode configuration
- HWU, LULCC possible
- Non local effects and remote feedbacks with large domain

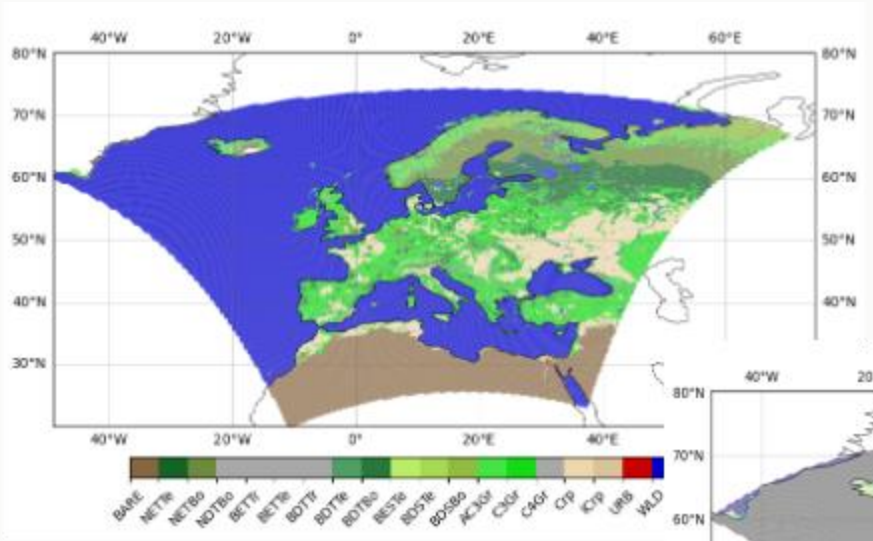
Grids:

icosahedral (ICON, eCLM)

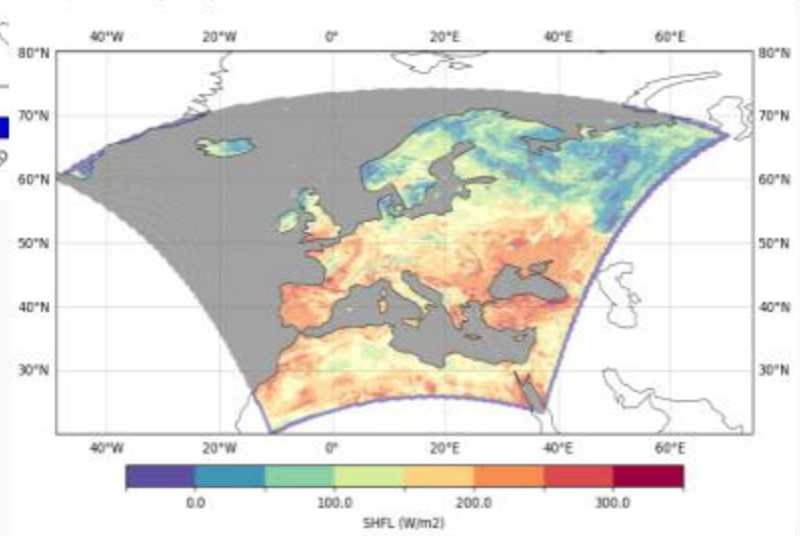
curvilinear (ParFlow)

4 different model grids:

- 12km domain
  - EUR-11; R13B05
- 3km domain
  - EUR-0275; R13B07



Dominant land use CORDEX EUR-11



EUR-11/R13B05 31th July 1990 12UTC

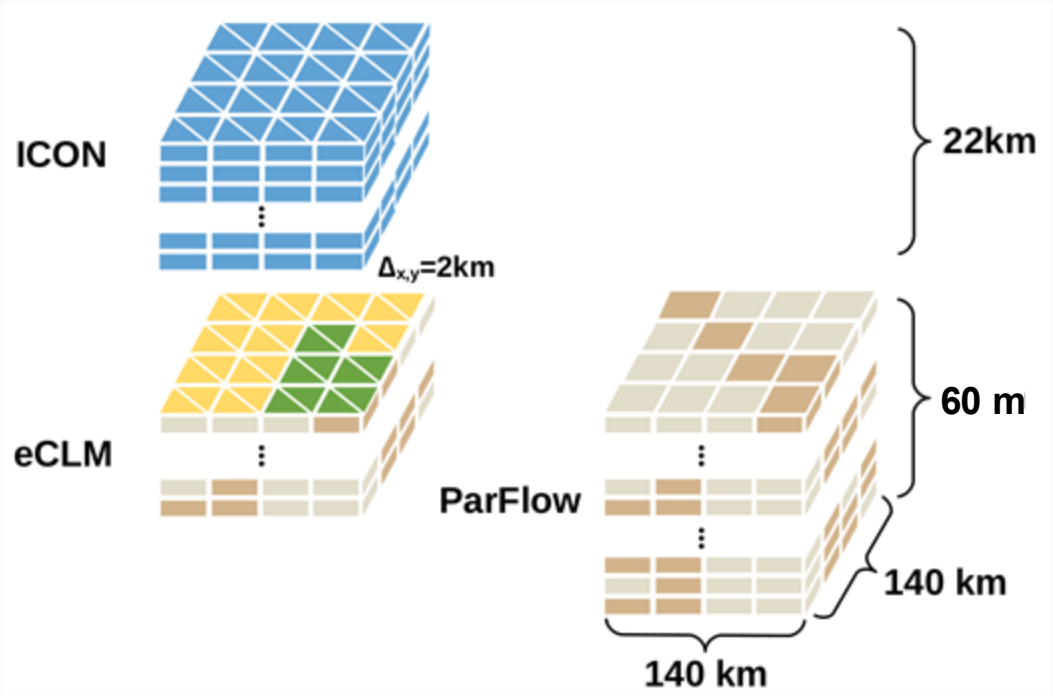
CLMcom, WG CPCS, CECPI-Ger initiative: optimized ICON configuration from coordinated community experiment

IDEALIZED TEST-CASE

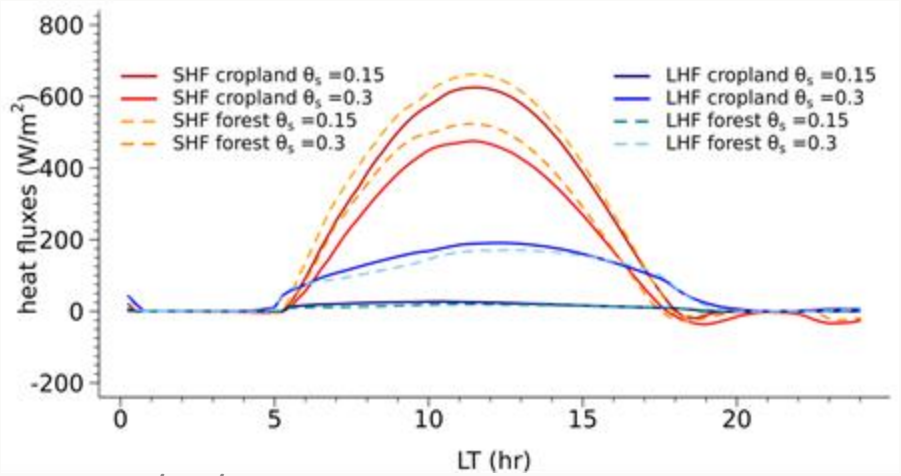
Sensitivity on land-cover and soil moisture

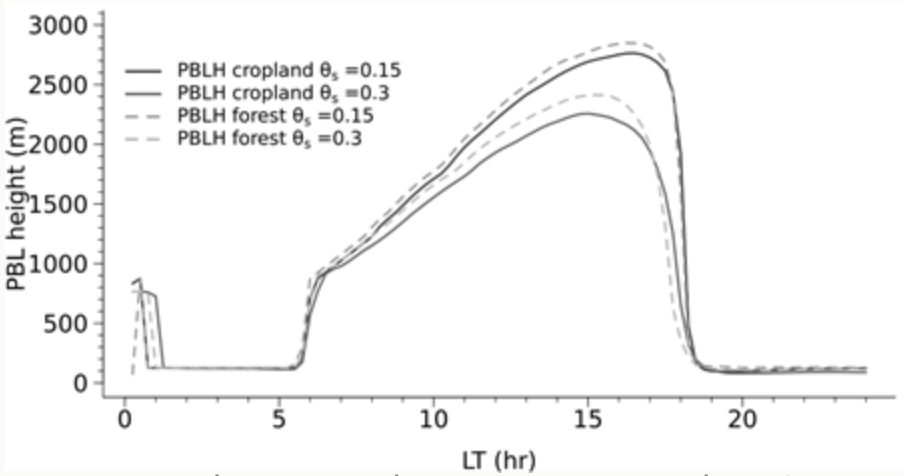
Setup

- ICON and eCLM on icosahedral grid, ParFlow on cartesian grid
- Homogeneous land and sub-surface with different initial conditions (24hrs)



Feedback on land-cover and soil moisture content:





Setup: double periodic boundary condition, atm. initialisation after Patton et al. (2005), loam soil

TSMP2 fully coupled simulation experiments, 3km frontier runs

Model setup and external parameter fields, 12km (R13B05, EUR-11), 3km (R13B07, EUR-0275)

Coupled G2A TSMP2

1600x1552x65/30

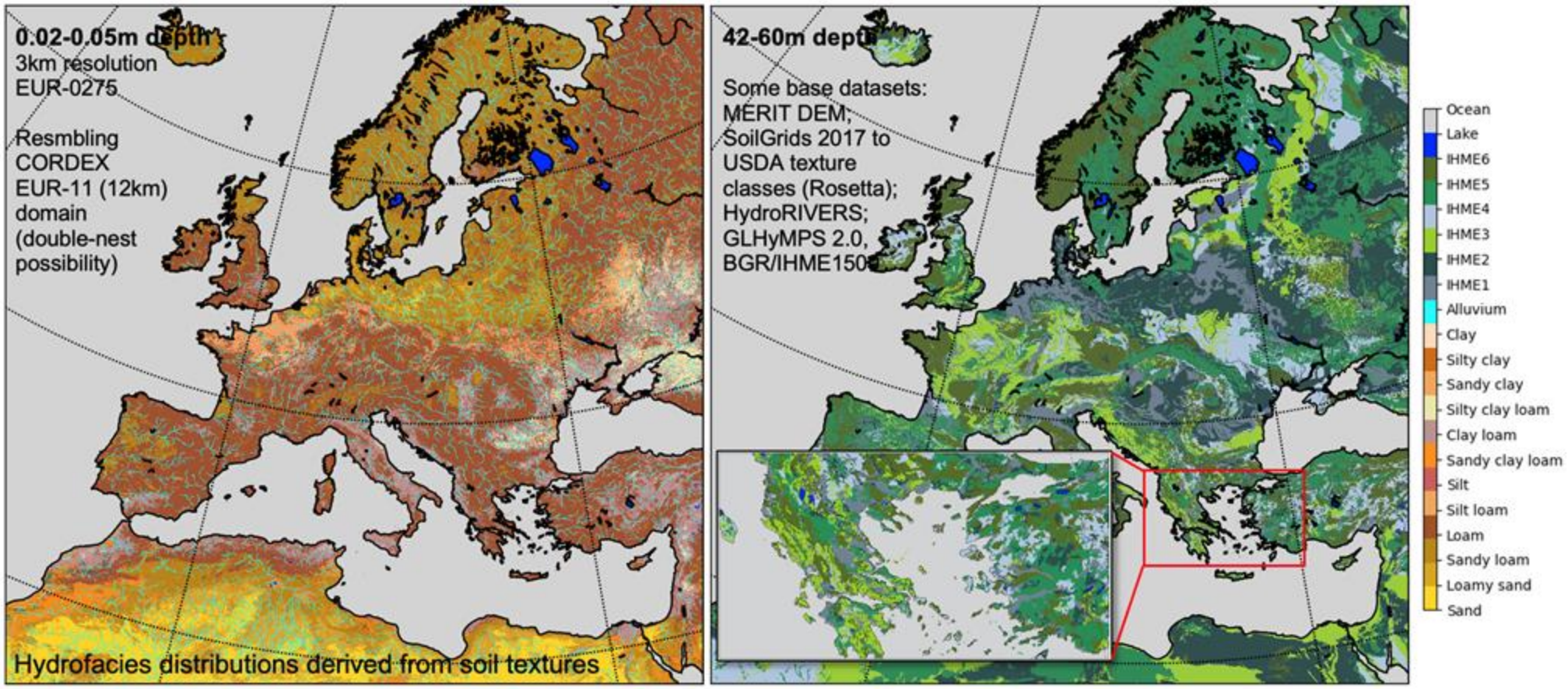
3km resolution

Grids: curvilinear (ParFlow), icosahedral (ICON, eCLM)

Climate mode configuration

HWU, LULCC possible

Non local effects and remote feedbacks with large domain



Goals | Build system | Work-flow engine | **Setups and Configuration** | Exa-scale readiness | Coming next

Additional tools and functionalities

TSMP2: Water Budget Analysis (WBA) and Data Tools

Goal: Evaluating model physical consistency using scalable data processing tools

Objective 1: Exploring Data Processing Tools

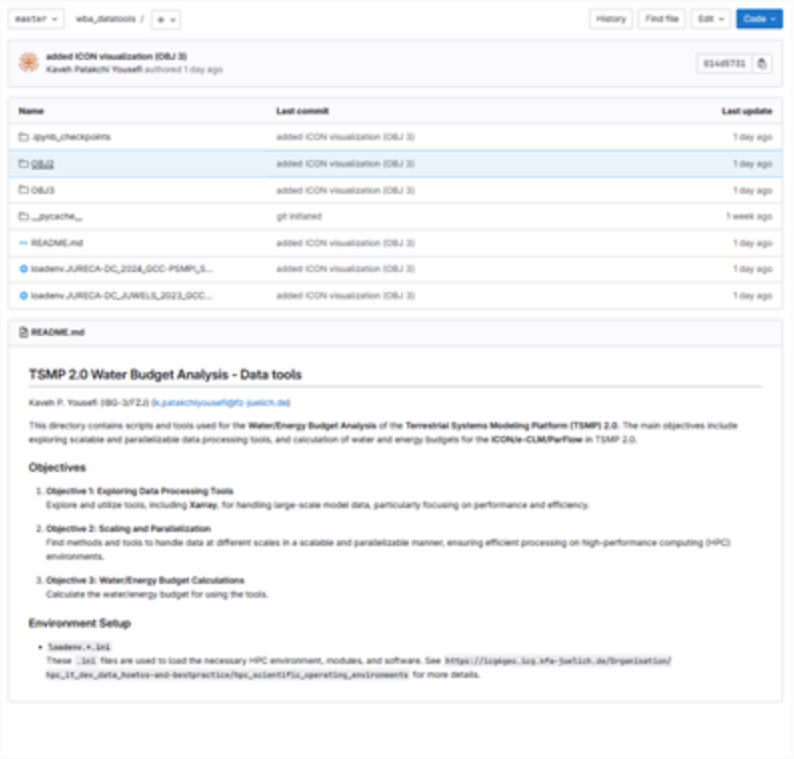
Explore and utilize tools, including Xarray, for handling large-scale model data, particularly focusing on performance and efficiency.

Objective 2: Scaling and Parallelization

Find methods and tools to handle data at different scales in a scalable and parallelizable manner, ensuring efficient processing on high-performance computing (HPC) environments.

Objective 3: Water Budget Calculations, physical consistency

Calculate the water budget for using the tools.



Gitlab repo for WBA and data tools (internal)

TSMP2-PDAF: Data Assimilation with TSMP2

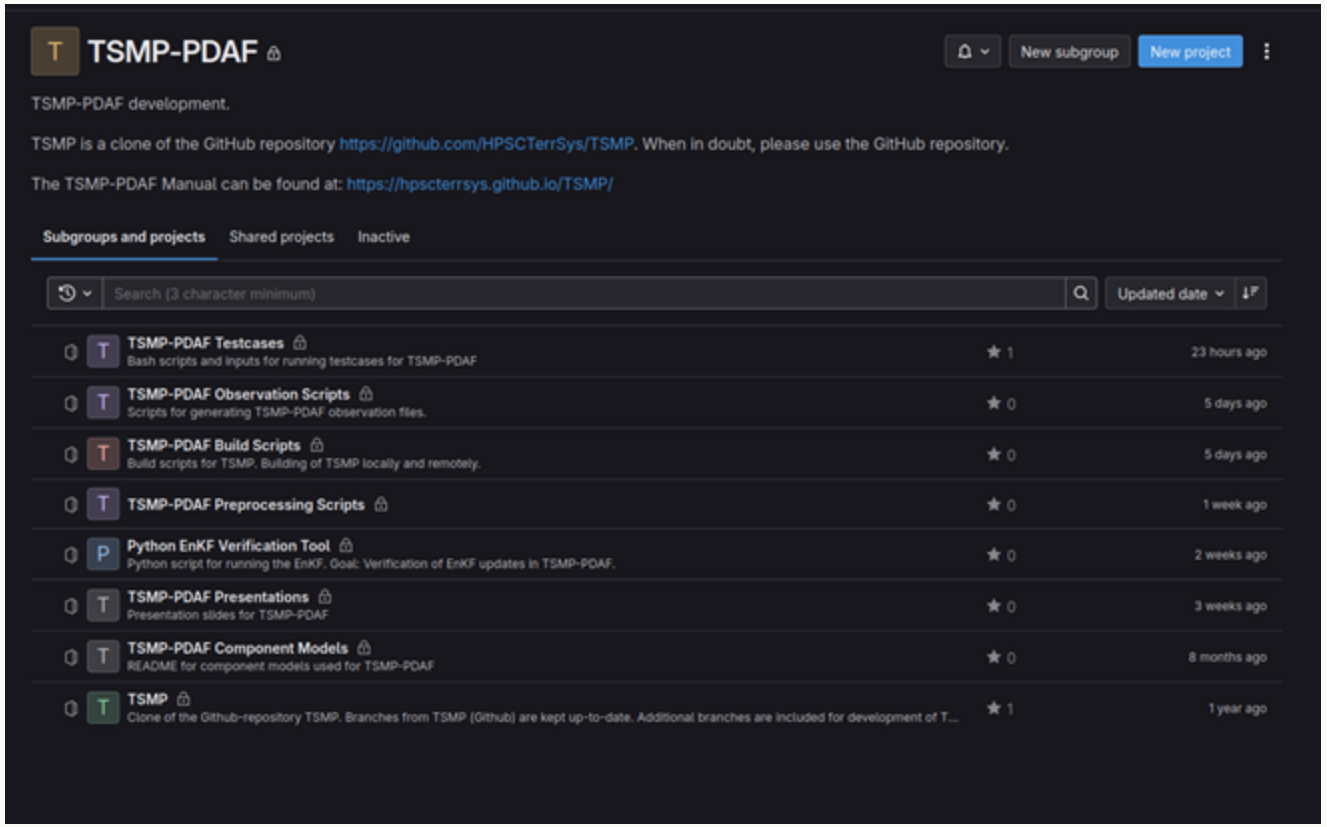
Collaborative development environment

Testing

- https://icg4geo.icg.kfa-juelich.de/ExternalRepos/tsmp-pdaf/TSMP-PDAF-testcases
- eCLM-PDAF testcases for SM-DA (10x10, NRW) and Snow-DA (single site)

Python

- https://icg4geo.icg.kfa-juelich.de/ExternalRepos/tsmp-pdaf
- Observation scripts
- Build scripts
- Preprocessing scripts (perturbation, ensemble namelist)
- Python verification tool



Goals | Build system | Work-flow engine | Setups and Configuration | Exa-scale readiness | Coming next

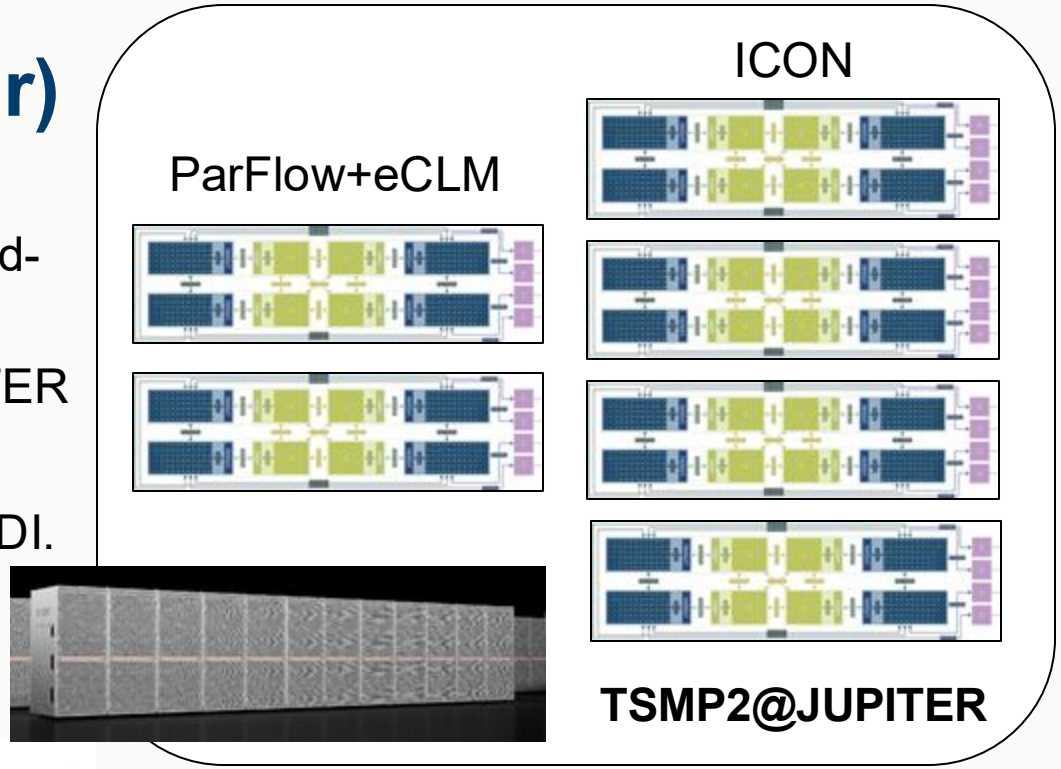
Preparing for exascale

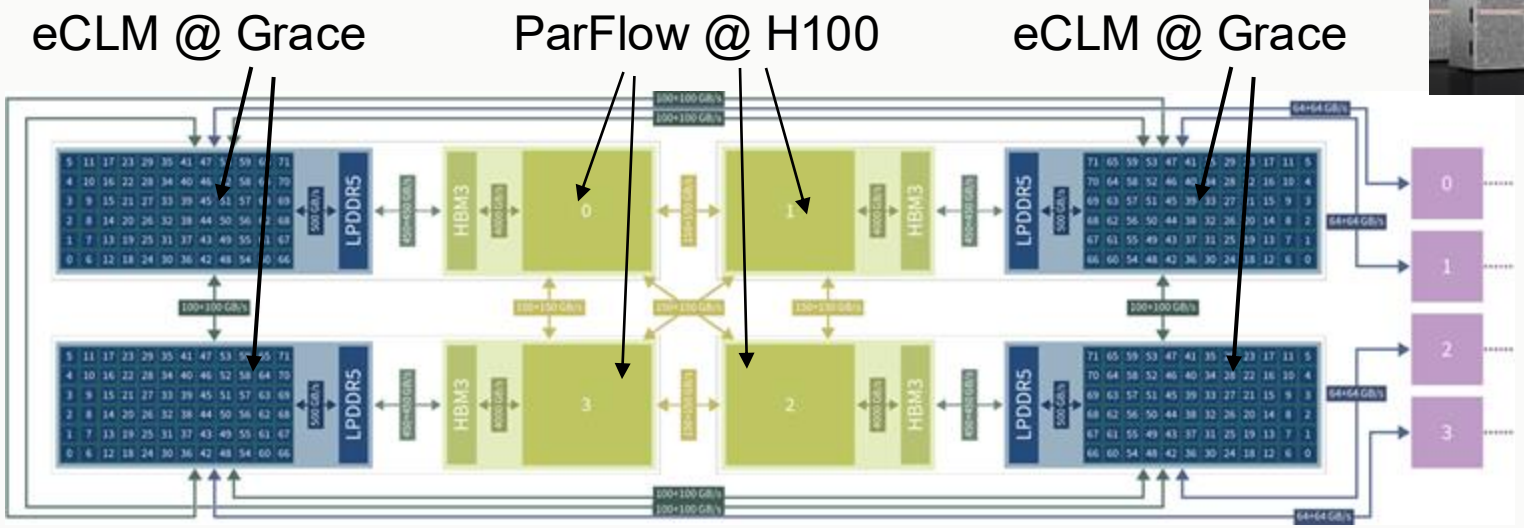
JUPITER will start production soon

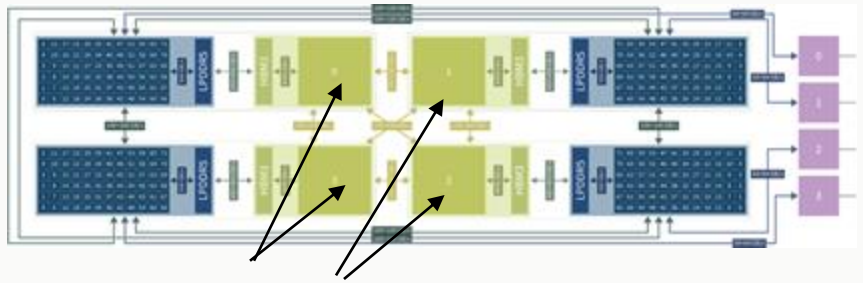
CRC1502
DETECT
https://sfb1502.de/

Towards TSMP2 in JUPITER(-Booster)

- ParFlow and ICON on GPUs, CLM on CPU
- Offloading the atmospheric model to GPUs should provide speed-up. Load balancing will completely change
- Exploit in-node heterogeneity in the GH200 superchips in JUPITER
- TSMP2 (CPU) running on JEDI
- First tests with eCLM (CPU) + ParFlow (GPU) are running in JEDI. Adding ICON (GPU) in progress.







In-node heterogeneous computing

JUPITER Booster Node: 4 x Nvidia Grace-Hopper superchip (GH200)
4 x Hopper H100 GPU + 4 x 72 Neoverse V2 ARM CPU cores

Goals

Build system

Work-flow engine

Setups and Configu ration

Exa-scale readine ss

Coming next

CONCLUSION & ONGOING WORK

Terrestrial Systems Modeling Platform (TSMP2) based on ICON, eCLM and ParFlow

Conclusion

- TSMP2 is an open-source, state-of-the-art RESM for comprehensive terrestrial water cycle research G2A
- Computational overhead of ~10-20%
  due to increase of physical representation
- Software infrastructure for fully coupled system and component models

Ongoing work

- Preparation of baseline simulation
  2000-2020 for Δ_x=12km and Δ_x=3km

TSMP2

https://github.com/HPSCTerrSys/TSMP2

Component models

ICON
https://gitlab.dkrz.de/icon/icon-model

eCLM
https://github.com/HPSCTerrSys/eCLM

ParFlow
https://github.com/parflow/parflow

Next steps

- Finish ongoing tests, tuning
- “TSMP2 introduction” and “TSMP2 frontier experiment” papers
- Start ERA5 reanalysis-driven 3km evaluation runs asap with latest CLMcom / CECPI-Ger configuration
- CI/CB tests for monitoring the system
- With the DETECT modelling community from 2025-07 onwards (help in improving, testing, tuning):
  - Use of build system and workflow engine (for all model combinations)
  - Getting accustomed to coupled TSMP2 in preparation for phase 2
- Make all data available via detectdata, also from short tests

CRC1502
DETECT
https://sfb1502.de/

Status of TSMP2 by DETECT/Z04

2025-05-14 I Stefan Poll[1,2,3], Marco van Hulten[4], Johannes Keller[1,2], Paul Rigor[1,2],
Daniel Caviedes-Voullième[2,3], Klaus Goergen[1,2]
Supported by Kaveh Patakchi Yousefi[1,2], Niklas Wager[1,2,*], Carl Hartick[1,2,*], Muhammad Fahad[1,2],
Ana Gonzalez-Nicolas[2,3]

[1]Institute of Bio- and Geosciences (IBG-3, Agrosphere), Research Centre Jülich (FZJ), Jülich, Germany
[2]Centre for High-Performance Scientific Computing in Terrestrial Systems (HPSC TerrSys), Geoverbund ABC/J, Jülich, Germany
[3]Simulation and Data Lab Terrestrial Systems (SDLTS), Jülich Supercomputing Centre (JSC), Research Centre Jülich (FZJ), Jülich, Germany
[4]High-Performance Computing and Analytics Lab (HPC/A-Lab), University of Bonn, Bonn, Germany
*New affiliation